

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:00:55 ; Search time 21.69 Seconds  
(without alignments)  
8.359 Million cell updates/sec

Title: US-09-214-836-1

Sequence: 1 KTWGQYNAV 9

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	91.4	9	1	US-08-787-547-61 Sequence 61, Appl
2	53	91.4	9	2	US-08-417-174-46 Sequence 46, Appl
3	53	91.4	9	2	US-08-902-516-29 Sequence 29, Appl
4	53	91.4	9	2	US-09-036-582-26 Sequence 26, Appl
5	53	91.4	10	2	US-08-417-174-47 Sequence 47, Appl
6	53	91.4	661	2	US-08-417-174-121 Sequence 121, Appl
7	53	91.4	668	1	US-07-891-942G-6 Sequence 6, Appl
8	50	86.2	661	2	US-08-417-174-27 Sequence 27, Appl
9	50	86.2	661	2	US-08-231-565A-27 Sequence 27, Appl
10	50	86.2	661	2	US-09-007-961-27 Sequence 27, Appl
11	48	82.8	9	2	US-08-417-174-71 Sequence 71, Appl
12	48	82.8	9	2	US-08-417-174-72 Sequence 72, Appl
13	48	82.8	9	2	US-08-417-174-73 Sequence 73, Appl
14	48	82.8	9	2	US-08-417-174-74 Sequence 74, Appl
15	48	82.8	9	2	US-08-417-174-75 Sequence 75, Appl
16	47	81.0	9	2	US-08-417-174-68 Sequence 68, Appl
17	47	81.0	9	2	US-08-417-174-69 Sequence 69, Appl
18	47	81.0	9	2	US-08-417-174-70 Sequence 70, Appl
19	44	75.9	9	2	US-08-417-174-76 Sequence 76, Appl
20	43	74.1	9	2	US-08-417-174-77 Sequence 77, Appl
21	43	74.1	9	2	US-08-417-174-78 Sequence 78, Appl
22	43	74.1	9	2	US-08-417-174-79 Sequence 79, Appl
23	43	74.1	9	2	US-08-417-174-80 Sequence 80, Appl
24	43	74.1	9	2	US-08-417-174-81 Sequence 81, Appl
25	43	74.1	9	2	US-08-417-174-82 Sequence 82, Appl
26	40	69.0	72	1	US-08-482-282B-6 Sequence 6, Appl
27	40	69.0	72	1	US-08-486-036A-6 Sequence 6, Appl

28	40	69.0	72	5	PCT-US96-09848-6	Sequence 6, Appl
29	40	69.0	401	5	PCT-US96-09848-31	Sequence 31, Appl
30	40	69.0	407	5	PCT-US96-09848-15	Sequence 15, Appl
31	39	67.2	585	1	US-08-212-188-4	Sequence 4, Appl
32	39	67.2	585	3	US-08-970-725-4	Sequence 4, Appl
33	39	67.2	585	5	PCT-US95-02708-4	Sequence 4, Appl
34	37	63.8	12	1	US-08-439-817-205	Sequence 205, App
35	37	63.8	20	1	US-08-241-054-85	Sequence 85, Appl
36	37	63.8	20	1	US-08-241-054-91	Sequence 91, Appl
37	37	63.8	20	1	US-08-390-156A-40	Sequence 40, Appl
38	37	63.8	20	1	US-08-390-156A-43	Sequence 43, Appl
39	37	63.8	20	1	US-08-439-817-65	Sequence 65, Appl
40	37	63.8	20	1	US-08-439-817-71	Sequence 71, Appl
41	37	63.8	20	1	US-08-485-508-85	Sequence 85, Appl
42	37	63.8	20	1	US-08-485-508-91	Sequence 91, Appl
43	37	63.8	191	1	US-08-044-621D-35	Sequence 35, Appl
44	37	63.8	191	1	US-08-709-912-11	Sequence 11, Appl
45	37	63.8	191	2	US-09-047-370-11	Sequence 11, Appl

#### ALIGNMENTS

RESULT 1  
US-08-787-547-61  
Sequence 61, Application US/08787547  
Patent No. 5783567  
GENERAL INFORMATION:  
APPLICANT: Hedley, Mary Lynne  
APPLICANT: Curley, Joanne M.  
APPLICANT: Langer, Robert S.  
TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY  
TITLE OF INVENTION: OF NUCLEIC ACID  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/787,547  
FILING DATE: 22-JAN-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-787-547-61  
Query Match 91.4%; Score 53; DB 1; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYWAY 9  
11111111  
DB 1 KTWGQYWAY 9

## RESULT 2

US-08-417-174-46

; Sequence 46, Application US/08417174

; Patent No. 5844075

## GENERAL INFORMATION:

APPLICANT: KAKAKAMI, YUTAKA; ROSENBERG,

APPLICANT: STEVEN A.

TITLE OF INVENTION: MELANOMA ANTIGENS AND

TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC

NUMBER OF SEQUENCES: 126

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN &amp; FINNEGAN, L.L.P.

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

## COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/417,174

FILING DATE: 05-APR-1995

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US/08/231,565

FILING DATE: 22-APR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CAROL M. GRUPPI

REGISTRATION NUMBER: 37,341

REFERENCE/DOCKET NUMBER: 2026-4124US1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

TELEX: 421792

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:

LENGTH: 9

TYPE: amino acid

STRANDEDNESS: Unknown

TOPOLOGY: Unknown

MOLECULE TYPE: Peptide

US-08-417-174-46

OY 1 KTWGQYWAY 9

11111111

DB 1 KTWGQYWAY 9

Query Match 91.4%; Score 53; DB 2; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## RESULT 3

US-08-902-516-29

; Sequence 29, Application US/08902516

; Patent No. 5891432

## GENERAL INFORMATION:

APPLICANT: SOO HOO, WILLIAM

TITLE OF INVENTION: MEMBRANE-BOUND CYTOKINE COMPOSITIONS

TITLE OF INVENTION: COMPRISING GM-CSF AND METHODS OF MODULATING AN IMMUNE

TITLE OF INVENTION: RESPONSE USING SAME

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:  
ADDRESSEE: CAMPBELL & FLORES, LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States  
ZIP: 92121

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/902,516

FILING DATE: 29-JUL-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-1M 2442

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619)535-9001

TELEFAX: (619)535-8949

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-902-516-29

Query Match 91.4%; Score 53; DB 2; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYWAY 9

11111111

DB 1 KTWGQYWAY 9

Query Match 91.4%; Score 53; DB 2; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## RESULT 4

US-09-036-582-26

; Sequence 26, Application US/09036582A

; Patent No. 5965381

## GENERAL INFORMATION:

APPLICANT: van der Bruggen, Pierre

TITLE OF INVENTION: DELIVERY OF PROTEINS INTO EUKARYOTIC CELLS

FILE REFERENCE: 11154

CURRENT APPLICATION NUMBER: US/09/036,582A

CURRENT FILING DATE: 1998-03-06

NUMBER OF SEQ ID NOS: 39

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 26

LENGTH: 9

TYPE: PRT

ORGANISM: Human gp100pme117 peptide

US-09-036-582-26

Query Match 91.4%; Score 53; DB 2; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYWAY 9

11111111

DB 1 KTWGQYWAY 9

Query Match 91.4%; Score 53; DB 2; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## RESULT 5

US-08-417-174-47

Sequence 47, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-47

Query Match 91.4%; Score 53; DB 2; Length 10;  
Best Local Similarity 88.9%; Pred. No. 0.012;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYNAV 9  
Db 1 KTWGQYNAV 9

RESULT 6  
US-08-417-174-121  
Sequence 121, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 121:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Protein  
US-08-417-174-121

Query Match 91.4%; Score 53; DB 2; Length 661;  
Best Local Similarity 88.9%; Pred. No. 0.75;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYNAV 9  
Db 154 KTWGQYNAV 162

RESULT 7  
US-07-891-942G-6  
Sequence 6, Application US/07891942G  
Patent No. 5679511  
GENERAL INFORMATION:  
APPLICANT: Kwon, Byoung Se  
TITLE OF INVENTION: cDNA CLONES FOR HUMAN TYROSINASE AND FOR  
TITLE OF INVENTION: A REGULATORY PROTEIN IN THE MELANIN PROTEIN PATHWAY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Christopher A. Michaels, Barnard, Brown &  
ADDRESSEE: Michaels  
STREET: 306 East State Street, Suite 220  
CITY: Ithaca  
STATE: NY  
COUNTRY: USA  
ZIP: 14850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/891,942G  
FILING DATE: 01-JUN-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 06/915,753  
FILING DATE: 06-OCT-1986  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/362,847  
FILING DATE: 07-JUN-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Michaels, Christopher A

REGISTRATION NUMBER: 34,390  
REFERENCE/DOCKET NUMBER: INDI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 607-273-1711  
TELEFAX: 607-273-2609  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 668 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PUBLICATION INFORMATION:  
AUTHORS: Kwon, Byoung Se  
AUTHORS: Chintamaneni, Chaya  
AUTHORS: Kozak, Christine A  
AUTHORS: Copeland, Neal G  
AUTHORS: Gilbert, Debra J  
AUTHORS: Jenkins, Nancy  
AUTHORS: Barton, David  
AUTHORS: Francke, Uta  
AUTHORS: Kobayashi, Yvonne  
AUTHORS: Kim, Kack K  
TITLE: A melanocyte-specific gene, Pmel 17, maps  
TITLE: near the silver coat color locus on mouse  
TITLE: chromosome 10 and is in a syntenic region on human  
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
VOLUME: 88  
PAGES: 9228-9232  
DATE: October 1991  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 668  
US-07-891-9426-6

Query Match  
Best Local Similarity 91.4%; Score 53; DB 1; Length 668;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYMW 9  
Db 154 KTWGQYMW 162

RESULT 8  
US-08-417-174-27  
Sequence 27, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAKAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-41240S1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Protein  
US-08-417-174-27

Query Match  
Best Local Similarity 86.2%; Score 50; DB 2; Length 661;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMW 7  
Db 154 KTWGQYMW 160

RESULT 9  
US-08-231-565A-27  
Sequence 27, Application US/08231565A  
Patent No. 5874560  
GENERAL INFORMATION:  
APPLICANT: KAKAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565A  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Protein  
US-08-231-565A-27



Query Match 86.2%; Score 50; DB 2; Length 661;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGOYW 7  
1111111  
Db 154 KTWGOYW 160

RESULT 10  
US-09-007-961-27  
; Sequence 27, Application US/09007961  
; Patent No. 5994523  
; GENERAL INFORMATION:  
; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
; APPLICANT: STEVEN A.  
; TITLE OF INVENTION: MELANOMA ANTIGENS AND  
; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
; TITLE OF INVENTION: METHODS  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/007,961  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/231,565  
; FILING DATE: 22-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAROL M. GRUPE  
; REGISTRATION NUMBER: 37,341  
; REFERENCE/DOCKET NUMBER: 2026-4124  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 661  
; TYPE: amino acid  
; STRANDEDNESS: Unknown  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: Protein  
US-09-007-961-27

Query Match 86.2%; Score 50; DB 2; Length 661;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGOYW 7  
1111111  
Db 154 KTWGOYW 160  
RESULT 11  
US-08-417-174-71  
; Sequence 71, Application US/08417174  
; Patent No. 5844075  
; GENERAL INFORMATION:  
; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
; APPLICANT: STEVEN A.

TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPE  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124051  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 71:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-71

Query Match 82.8%; Score 48; DB 2; Length 9;  
Best Local Similarity 87.5%; Pred. No. 1.4e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TWGOYMW 9  
1111111  
Db 2 TWGOYMW 9

RESULT 12  
US-08-417-174-72  
; Sequence 72, Application US/08417174  
; Patent No. 5844075  
; GENERAL INFORMATION:  
; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
; APPLICANT: STEVEN A.  
; TITLE OF INVENTION: MELANOMA ANTIGENS AND  
; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
; TITLE OF INVENTION: METHODS  
; NUMBER OF SEQUENCES: 126  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 72:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-72

Query Match 82.8%; Score 48; DB 2; Length 9;  
Best Local Similarity 87.5%; Pred. No. 1.4e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TWGQYNAV 9  
DB 2 TWGQYMOV 9

RESULT 13  
US-08-417-174-73  
Sequence 73, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 73:

SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-73

Query Match 82.8%; Score 48; DB 2; Length 9;  
Best Local Similarity 87.5%; Pred. No. 1.4e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TWGQYNAV 9  
DB 2 TWGQYMOV 9

RESULT 14  
US-08-417-174-74  
Sequence 74, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 74:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-74

Query Match 82.8%; Score 48; DB 2; Length 9;  
Best Local Similarity 87.5%; Pred. No. 1.4e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TWGQYNAV 9  
DB 2 TWGQYMOV 9

RESULT 15  
US-08-417-174-75  
Sequence 75, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 75:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-75

Query Match 82.8%; Score 48; DB 2; Length 9;  
Best Local Similarity 87.5%; Pred. No. 1.4e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TWGQYWAY 9  
Db 2 TWGQYWAY 9

RESULT 16  
US-08-417-174-68  
Sequence 68, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE

CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-68

Query Match 81.0%; Score 47; DB 2; Length 9;  
Best Local Similarity 77.8%; Pred. No. 1.4e+05;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTWGQYWAY 9  
Db 1 KTWGQYWAY 9

RESULT 17  
US-08-417-174-69  
Sequence 69, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-69

Query Match 81.0%; Score 47; DB 2; Length 9;  
Best Local Similarity 77.8%; Pred. No. 1.4e+05;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTWGQYWAY 9  
| | | | | | |  
Db 1 KTWGQYWAY 9

RESULT 18  
US-08-417-174-70  
Sequence 70, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAKAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A. MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-70

Query Match 81.0%; Score 47; DB 2; Length 9;  
Best Local Similarity 77.8%; Pred. No. 1.4e+05;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTWGQYWAY 9  
| | | | | | |  
Db 1 KTWGQYWAY 9

RESULT 19  
US-08-417-174-76  
Sequence 76, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAKAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A. MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-76

Query Match 75.9%; Score 44; DB 2; Length 9;  
Best Local Similarity 77.8%; Pred. No. 1.4e+05;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYWAY 9  
| | | | | | |  
Db 1 KTWGQYWAY 9

RESULT 20  
US-08-417-174-77  
Sequence 77, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:

APPLICANT: KAKAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124051  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 77:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-77

Query Match 74.1%; Score 43; DB 2; Length 9;  
Best Local Similarity 77.8%; Pred. No. 1.4e+05;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYNAV 9  
||:|||||  
Db 1 KTFQGYMOV 9

RESULT 21  
US-08-417-174-78  
Sequence 78, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAKAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124051  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-78

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124051  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-78

Query Match 74.1%; Score 43; DB 2; Length 9;  
Best Local Similarity 85.7%; Pred. No. 1.4e+05;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WGYWNAV 9  
|||||  
Db 3 WGYWMOV 9

RESULT 22  
US-08-417-174-79  
Sequence 79, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAKAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124051  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-78

TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 79:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS: Unknown  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: Peptide  
US-08-417-174-79

Query Match 74.1%; Score 43; DB 2; Length 9;  
Best Local Similarity 85.7%; Pred. No. 1.4e+05;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WGOYMAV 9  
Db 3 WGOYMOV 9

RESULT 23  
US-08-417-174-80  
; Sequence 80, Application US/08417174  
; Patent No. 5844075  
; GENERAL INFORMATION:  
; APPLICANT: KAKAKAMI, YUTAKA; ROSENBERG,  
; APPLICANT: STEVEN A.  
; TITLE OF INVENTION: MELANOMA ANTIGENS AND  
; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
; TITLE OF INVENTION: METHODS  
; NUMBER OF SEQUENCES: 126  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/417,174  
; FILING DATE: 05-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/231,565  
; FILING DATE: 22-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAROL M. GRUPPI  
; REGISTRATION NUMBER: 37,341  
; REFERENCE/DOCKET NUMBER: 2026-4124US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 80:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS: Unknown  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: Peptide  
US-08-417-174-80

Query Match 74.1%; Score 43; DB 2; Length 9;  
Best Local Similarity 85.7%; Pred. No. 1.4e+05;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WGOYMAV 9

Db 3 WGOYMOV 9

RESULT 24  
US-08-417-174-81  
; Sequence 81, Application US/08417174  
; Patent No. 5844075  
; GENERAL INFORMATION:  
; APPLICANT: KAKAKAMI, YUTAKA; ROSENBERG,  
; APPLICANT: STEVEN A.  
; TITLE OF INVENTION: MELANOMA ANTIGENS AND  
; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
; TITLE OF INVENTION: METHODS  
; NUMBER OF SEQUENCES: 126  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/417,174  
; FILING DATE: 05-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/231,565  
; FILING DATE: 22-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAROL M. GRUPPI  
; REGISTRATION NUMBER: 37,341  
; REFERENCE/DOCKET NUMBER: 2026-4124US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 81:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS: Unknown  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: Peptide  
US-08-417-174-81

Query Match 74.1%; Score 43; DB 2; Length 9;  
Best Local Similarity 85.7%; Pred. No. 1.4e+05;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WGOYMAV 9  
Db 3 WGOYMOV 9

RESULT 25  
US-08-417-174-82  
; Sequence 82, Application US/08417174  
; Patent No. 5844075  
; GENERAL INFORMATION:  
; APPLICANT: KAKAKAMI, YUTAKA; ROSENBERG,  
; APPLICANT: STEVEN A.  
; TITLE OF INVENTION: MELANOMA ANTIGENS AND  
; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
; TITLE OF INVENTION: METHODS  
; NUMBER OF SEQUENCES: 126  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 82:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-82

Query Match 74.1%; Score 43; DB 2; Length 9;  
Best Local Similarity 85.7%; Pred. No. 1.4e+05;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 WGYWAV 9  
Db 3 WGYWAV 9

RESULT 26  
US-08-482-282B-6  
Sequence 6, Application US/08482282B  
Patent No. 5792624  
GENERAL INFORMATION:  
APPLICANT: TRIPP, Cynthia A.  
APPLICANT: Wisniewski, Nancy  
APPLICANT: Gileve, Robert B.  
APPLICANT: Frank, Glenn R.  
TITLE OF INVENTION: NOVEL FILARIID NEMATODE CYSTEINE  
TITLE OF INVENTION: PROTEASE PROTEIN, NUCLEIC ACID MOLECULES AND USES  
TITLE OF INVENTION: THEROP  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80525  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Wordperfect for Windows, Version 7.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,282B

FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: 2618-33-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (970) 493-7272  
TELEFAX: (970) 484-9505  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-482-282B-6

Query Match 69.0%; Score 40; DB 1; Length 72;  
Best Local Similarity 55.6%; Pred. No. 7.5;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGYWAV 9  
Db 30 KTYGEYWT 38

RESULT 27  
US-08-486-036A-6  
Sequence 6, Application US/08486036A  
Patent No. 5795768  
GENERAL INFORMATION:  
APPLICANT: TRIPP, Cynthia A.  
APPLICANT: Wisniewski, Nancy  
APPLICANT: Gileve, Robert B.  
APPLICANT: Frank, Glenn R.  
TITLE OF INVENTION: NOVEL FILARIID NEMATODE  
TITLE OF INVENTION: CYSTEINE PROTEASE PROTEINS, NUCLEIC ACID  
TITLE OF INVENTION: MOLECULES AND USES THEREOF  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross & McIntosh  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,036A  
FILING DATE: June 7, 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-33  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-486-036A-6

Query Match 69.0%; Score 40; DB 1; Length 72;

Best Local Similarity 55.6%; Pred. No. 7.5;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYWAY 9  
11:1111:  
DB 30 KTYGEYWI 38

## RESULT 28

PCT-US96-09848-6  
; Sequence 6, Application PC/TUS9609848  
; GENERAL INFORMATION:  
; APPLICANT: Tripp, Cynthia A.  
; APPLICANT: Wisniewski, Nancy  
; TITLE OF INVENTION: NOVEL FILARIID NEMATODE CYSTEINE  
; TITLE OF INVENTION: PROTEASE PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross & McIntosh  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: U.S.A.  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/09848  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/486,036  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-33-PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 72 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US96-09848-6

Query Match 69.0%; Score 40; DB 5; Length 72;  
Best Local Similarity 55.6%; Pred. No. 7.5;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYWAY 9  
11:1111:  
DB 30 KTYGEYWI 38

## RESULT 29

PCT-US96-09848-31  
; Sequence 31, Application PC/TUS9609848  
; GENERAL INFORMATION:  
; APPLICANT: Tripp, Cynthia A.  
; APPLICANT: Wisniewski, Nancy  
; TITLE OF INVENTION: NOVEL FILARIID NEMATODE CYSTEINE  
; TITLE OF INVENTION: PROTEASE PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross & McIntosh  
; STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80203

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/09848  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/486,036  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-33-PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 401 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-09848-31

Query Match 69.0%; Score 40; DB 5; Length 401;  
Best Local Similarity 55.6%; Pred. No. 41;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYWAY 9  
11:1111:  
DB 359 KTYGEYWI 367

## RESULT 30

PCT-US96-09848-15  
; Sequence 15, Application PC/TUS9609848  
; GENERAL INFORMATION:  
; APPLICANT: Tripp, Cynthia A.  
; APPLICANT: Wisniewski, Nancy  
; TITLE OF INVENTION: NOVEL FILARIID NEMATODE CYSTEINE  
; TITLE OF INVENTION: PROTEASE PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross & McIntosh  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: U.S.A.  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/09848  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/486,036  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-33-PCT



TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 407 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-09848-15

Query Match 69.0%; Score 40; DB 5; Length 407;  
Best Local Similarity 55.6%; Pred. No. 42;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KTWGQYMAV 9  
||:|:|:  
Db 365 KTYGEYWI 373

Search completed: July 3, 2001, 10:22:35  
Job time: 1300 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:22:40 ; Search time 37.77 Seconds  
(without alignments)  
31.526 Million cell updates/sec

Title: US-09-214-836-1  
Perfect score: 58  
Sequence: 1 KTWGQYMAV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues  
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: SP-archaea:\*  
2: SP-bacteria:\*  
3: SP-fungi:\*  
4: SP-human:\*  
5: SP-invertebrate:\*  
6: SP-mammal:\*  
7: SP-mhc:\*  
8: SP-organelle:\*  
9: SP-phage:\*  
10: SP-plant:\*  
11: SP-protent:\*  
12: SP-unclassified:\*  
13: SP-vertebrate:\*  
14: SP-virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	72.4	902	2	Q9S1E4
2	41	70.7	116	3	Q9UVZ4
3	41	70.7	458	5	Q18533
4	40	69.0	276	1	Q9Y9N5
5	40	69.0	281	2	Q9KPY3
6	40	69.0	305	2	Q45818
7	40	69.0	400	2	P74474
8	40	69.0	518	5	O02622
9	40	69.0	750	2	Q9REPI
10	40	69.0	1451	5	O01737
11	39	67.2	73	14	Q9DMY8
12	39	67.2	191	2	Q9EW89
13	39	67.2	228	2	O59962
14	39	67.2	236	2	Q9RXG5
15	39	67.2	259	2	P70874
16	39	67.2	281	11	O55126
17	39	67.2	285	4	O43801
18	39	67.2	286	4	O75323
19	39	67.2	288	13	Q9PU58

20	39	67.2	335	2	O08346
21	39	67.2	335	2	Q9RM4
22	39	67.2	358	13	Q9W616
23	39	67.2	358	13	Q9PW55
24	39	67.2	362	13	Q9W617
25	39	67.2	366	13	Q9DFC6
26	39	67.2	367	13	Q9DFC5
27	39	67.2	367	13	Q9PD36
28	39	67.2	469	1	O50200
29	39	67.2	545	10	Q9ZPS2
30	39	67.2	559	10	Q9SDN4
31	39	67.2	580	10	O04585
32	39	67.2	580	10	O9ZTX4
33	39	67.2	1081	2	Q9PMS7
34	38	65.5	73	14	Q9DMY4
35	38	65.5	209	2	P73834
36	38	65.5	255	10	Q9FL76
37	38	65.5	295	5	Q9U2E7
38	38	65.5	310	2	O45279
39	38	65.5	317	2	P97024
40	38	65.5	320	2	Q9F4V6
41	38	65.5	329	2	Q9RMH9
42	38	65.5	336	5	Q9VNR6
43	38	65.5	498	2	O87798
44	38	65.5	501	2	O07068
45	38	65.5	569	5	O62270

# ALIGNMENTS

RESULT 1  
ID Q9S1E4 PRELIMINARY; PRT; 902 AA.  
AC Q9S1E4;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE NREI PROTEIN.  
GN NREI.  
OS Wolinella succinogenes.  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Wolinella.  
OX NCBI\_Taxid-844;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Simon J., Gross R., Einsle O., Kroneck P.M.H., Kroeger A., Klimek O.;  
RT "A NapC/NitF-type cytochrome c (NirH) is the mediator between the  
RT quinone pool and the cytochrome c nitrite reductase of Wolinella  
RT succinogenes.";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ245540; CAB53161.1; -;  
DR InterPro; IPR001064; -;  
DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; UNKNOWN.1.  
SQ SEQUENCE 902 AA; 102016 MW; D2621BF04288380 CRC64;

Query Match 72.4%; Score 42; DB 2; Length 902;  
Best Local Similarity 62.5%; Pred. No. 74;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGQYMA 8  
DB 794 ESMGRYMA 801  
RESULT 2  
ID Q9UVZ4 PRELIMINARY; PRT; 116 AA.  
AC Q9UVZ4;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE XYLANASE (FRAGMENT).  
GN AXYL2.  
OS Gaumannomyces graminis.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetes Incertae sedis; Magnaporthaceae; Gaumannomyces.  
OX NCBI\_TaxID=29850;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Emami K.;  
RT "PCR-based characterization of fungal xylanase genes."  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ249160; CAB53513.1; -.  
DR HSSP; P09850; 1XNB.  
DR InterPro; IPR001137; -.  
DR Pfam; PF00457; Glyco\_hydro.11; 1.  
DR PRINTS; PR00911; GLYHDBLASE1.  
DR PROSITE; PS00776; GLYCOSYL\_HYDROL\_F11\_1; 1.  
FT NON\_TER 1  
FT NON\_TER 116  
SQ SEQUENCE 116 AA; 12791 MM; DA7BA4FEA6770E9E CRC64;

Query Match 70.7%; Score 41; DB 3; Length 116;  
Best Local Similarity 66.7%; Pred. No. 13;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RTWQGYWAV 9  
II: IIII:  
DB 80 RTWQGYWAV 88

RESULT 3  
ID 018533 PRELIMINARY; PRT; 458 AA.  
AC 018533:  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE PREPROCATHEPSIN C PRECURSOR (EC 3.4.14.1).  
OS Schistosoma japonicum (Blood fluke).  
OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;  
OC Trematoda; Digenea; Strigoida; Schistosomatidae; Schistosomatidae;  
OC Schistosoma.  
OX NCBI\_TaxID=6182;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CHINESE;  
RX MEDLINE=97442731; PubMed=9297696;  
RA Brindley P.J., Kallina B.H., Dalton J.P., Day S.R., Wong J.Y.,  
RA Smythe M.L., Mcmanus D.P.;  
RT "Proteolytic degradation of host hemoglobin by schistosomes."  
RL Mol. Biochem. Parasitol. 89:1-9(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CHINESE;  
RA Hota-Jamriska L., Dalton J.P., Askov J., Day S.R., Fan J.,  
RA Brindley P.J.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U77932; AAC32040.1; -.  
DR HSSP; P00787; THE.  
DR InterPro; IPR000169; -.  
DR InterPro; IPR000668; -.  
DR Pfam; PF00112; Peptidase\_C1; 1.  
DR PRINTS; PS00705; PAPA1N.  
DR PROSITE; PS00139; THIOL\_PROTEASE\_CYS; 1.  
DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
DR Hydrolase; Signal; Thiol protease.  
FT SIGNAL 1  
FT CHAIN 222 458  
SQ SEQUENCE 458 AA; 52698 MM; ADA976566C4142C CRC64;

Query Match 70.7%; Score 41; DB 5; Length 458;

Best Local Similarity 83.3%; Pred. No. 54;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 TWGQY 7  
II: IIII:  
DB 425 SWGQY 430

RESULT 4  
ID 0919N5 PRELIMINARY; PRT; 276 AA.  
AC 0919N5:  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE 276AA LONG HYPOTHETICAL LACTOSE TRANSPORT SYSTEM PERMEASE PROTEIN.  
GN APER2253.  
OS Aeropyrum pernix.  
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;  
OC Aeropyrum.  
OX NCBI\_TaxID=56636;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K1.  
RX MEDLINE=99310339; PubMed=10382966;  
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,  
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,  
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
RT "Complete genome sequence of an aerobic hyper-thermophilic  
crenarchaeon, Aeropyrum pernix K1."  
RL DNA Res. 6:83-101(1999).  
DR EMBL; AP000064; BAA81265.1; -.  
DR InterPro; IPR000515; -.  
DR Pfam; PF00528; BPD\_transp; 1.  
SQ SEQUENCE 276 AA; 31573 MM; A94B6CBF8032631 CRC64;

Query Match 69.0%; Score 40; DB 1; Length 276;  
Best Local Similarity 55.6%; Pred. No. 46;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 RTWQGYWAV 9  
II: IIII:  
DB 215 RTWQGYWAV 223

RESULT 5  
ID 09KPY3 PRELIMINARY; PRT; 281 AA.  
AC 09KPY3:  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE HYPOTHETICAL PROTEIN VC2229.  
GN VC2229.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-EL TOR N16961 / SEROTYPE O1;  
RX MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva L.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Uitterlinden T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio

RT cholerae.";  
RL Nature 406:477-483(2000).  
DR EMBL: AE004294; AAF55373.1;  
DR TIGR: VC2229;  
SQ SEQUENCE 281 AA; 31824 MW; 7CA75AD3494DEFDD0 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 40; DB 2; Length 281;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGOYX 7  
|||||  
DB 99 WGOYX 103

RESULT 6  
ID 045818 PRELIMINARY; PRT; 305 AA.  
AC 045818;  
DT 01-NOV-1996 (TREMblrel. 01, Created)  
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
DE HYPOTHETICAL 33.7 KDA PROTEIN (BGHC).  
GN BGHC.  
OS Chloroflexus aurantiacus.  
OC Bacteria; Green non-sulfur bacteria; Chloroflexaceae group;  
OC Chloroflexaceae; Chloroflexus.  
OX NCBI\_TaxID=1108;  
RN [1]  
RP SEQUENCE OF 224-305 FROM N.A.  
RC STRAIN-J10-FL;  
RX MEDLINE=94192803; PubMed=7511541;  
RA Nidemeier G., Shiozawa J., Lottspeich F., Feick R.;  
RT "The primary structure of two chlorosome proteins from Chloroflexus  
aurantiacus.";  
RL FEBS Lett. 342:61-65(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-J10-FL;  
RA Nidemeier G.;  
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94192803; PubMed=7511541;  
RA Nidemeier G., Shiozawa J.A., Lottspeich F., Feick R.G.;  
RT "The primary structure of two chlorosome proteins from Chloroflexus  
aurantiacus.";  
RL FEBS Lett. 342:61-65(1994).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20433268; PubMed=10976061;  
RA Xiong J., Fischer W.M., Inoue K., Nakahara M., Bauer C.E.;  
RT "Molecular evidence for the early evolution of photosynthesis.";  
RL Science 289:1724-1730(2000).  
DR EMBL: Z34000; CAA83969.1;  
DR EMBL: AF288602; AAG15233.1;  
DR InterPro: IPR000537;  
DR Pfam: PF01040; UblA; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 305 AA; 33674 MW; F990F92F2D1C2B07 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 40; DB 2; Length 305;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGOYX 7  
|||||  
DB 246 WGOYX 250

RESULT 7

P74474  
ID P74474 PRELIMINARY; PRT; 400 AA.  
AC P74474;  
DT 01-FEB-1997 (TREMblrel. 02, Created)  
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)  
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
DE D-ALANYL-D-ALANINE CARBOXYPEPTIDASE.  
GN SUR1924.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirosewa M., Sugiyama M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
Synechocystis sp. strain PCC6803. II. Sequence determination of the  
entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
DR EMBL: D90915; BAA18575.1;  
DR MEROPS: S12.001;  
KW Carboxypeptidase.  
SQ SEQUENCE 400 AA; 44316 MW; 75510481820E462F CRC64;

Query Match  
Best Local Similarity 100.0%; Score 40; DB 2; Length 400;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGOYX 7  
|||||  
DB 335 WGOYX 339

RESULT 8  
ID 002622 PRELIMINARY; PRT; 518 AA.  
AC 002622;  
DT 01-JUL-1997 (TREMblrel. 04, Created)  
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)  
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1) (1/4-ALPHA-D-GLUCAN  
DE GLUCANOHYDROLASE) (GLYCOGENASE) (FRAGMENT).  
GN AMY.  
OS Crassostrea gigas (Pacific oyster).  
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Ostreoida;  
OC Ostreoida; Ostreidae; Crassostrea.  
OX NCBI\_TaxID=29159;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-DIGESTIVE GLAND;  
RA Moal J., Daniel J.Y., Le Moine S., Sello D., Van Wormhoudt A.,  
RA Samain J.F.;  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
CC -/- CATALYTIC ACTIVITY: ENDOMYDOLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
DR EMBL: Y08370; CAA69658.1;  
DR HSP: P00690; 1FH.  
DR InterPro: IPR000461;  
DR Pfam: PF00128; alpha-amylase; 1.  
DR PRINTS: PR00110; ALPHAAMYLASE.  
KW Signal; Hydrolase; Glycosidase.  
FT NON\_TER 1  
FT SIGNAL <1 18 POTENTIAL.  
FT CHAIN 19 518 ALPHA-AMYLASE.  
SQ SEQUENCE 518 AA; 57435 MW; 2F3864914E77278AC CRC64;

Query Match  
Best Local Similarity 100.0%; Score 40; DB 5; Length 518;

Best Local Similarity 55.6%; Pred. No. 87;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTMGOYMA 9  
||||:|  
DB 292 KTMGOYMA 300

RESULT 9

O9REPI PRELIMINARY; PRT; 750 AA.

AC O9REPI 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE FERRECHROME RECEPTOR FCUA PRECURSOR.  
GN FCUA.  
OS Zymomonas mobilis.  
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonas group;  
OX NCBI\_TaxID=542;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ZM4;  
RA Um H.W., Kang H.S.;  
RI "Sequence analysis of 42B3 fosmid clone of Zymomonas mobilis."  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF213822; AAF23804.1; -  
KW Receptor.  
SQ SEQUENCE 750 AA; 82315 MW; 272729049A69FCED CRC64;

Query Match 69.0%; Score 40; DB 2; Length 750;  
Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTMGOYMA 8  
||||:|  
DB 313 KTMGOYMA 320

RESULT 10

O01737 PRELIMINARY; PRT; 1451 AA.

AC O01737 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE F20H11.2 PROTEIN.  
GN F20H11.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).

RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Wilson R., Moesner J., Graves T.;  
RT "The sequence of C. elegans cosmid F20H11."  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Waterston R.;  
KW Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF002197; AAD34660.1; -  
SQ SEQUENCE 1451 AA; 161680 MW; BCC3557BD68F8E01 CRC64;

Query Match 69.0%; Score 40; DB 5; Length 1451;  
Best Local Similarity 83.3%; Pred. No. 2.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOYMA 8  
||||:|  
DB 572 WGOYMA 577

RESULT 11

O9DMY8 PRELIMINARY; PRT; 73 AA.

AC O9DMY8 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE MAJOR CAPSID PROTEIN L1 (FRAGMENT).  
OS Human papillomavirus.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OX NCBI\_TaxID=10566;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-FA50;  
RA Antonsson A., Hossain S., Simon M., Hansson B.G.;  
RT "Skin HPV found on foreheads."  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY009882; AAG27044.1; -  
FT NON\_TER  
FT 1 73  
SQ SEQUENCE 73 AA; 8479 MW; F1DSBC6307D39062 CRC64;

Query Match 67.2%; Score 39; DB 14; Length 73;  
Best Local Similarity 71.4%; Pred. No. 17;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 WGOYMA 9  
||||:|  
DB 42 WGOYMA 48

RESULT 12

O9EW89 PRELIMINARY; PRT; 191 AA.

AC O9EW89 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE BETA-XYLANASE (EC 3.2.1.8) (FRAGMENT).  
GN GXYN1.  
OS Streptomyces olivaceoviridis (Streptomyces corchorus11).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1921;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-E-86;  
RA Bin Y.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-E-86;  
RA Zhang H., Yao B., Wang Y.;  
RT "Cloning and expression of the 23kD beta-xylanase gene from  
Streptomyces olivaceoviridis E-86."  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ292317; CAC19491.1; -  
KW Xylan degradation; Hydrolase; Glycosidase.

FT NON TER 1 1  
SQ SEQUENCE 191 AA; 20781 MW; 7E1C45AADE1B6B9C CRC64;

Query Match  
Best Local Similarity 67.2%; Score 39; DB 2; Length 191;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTMGOYNAV 9  
DB 131 KTFNOYMSV 139

RESULT 13  
Q59962 PRELIMINARY; PRT; 228 AA.

AC Q59962;  
DT 01-NOV-1996 (TREMblrel. 01, Created)  
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
DE ENDO-1,4-BETA-XYLANASE (EC 3.2.1.8) (1,4-BETA-D-XYLAN  
DE XYLANOXYDROLASE).  
GN XYL.  
OS Streptomyces sp.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1931;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=538;  
RA Georjia J.C.E.;  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC  
CC LINKAGES IN XYLANS.  
DR EMBL; X9518; CA67143.1; -  
DR HSSP; P09850; 1XNB.  
DR InterPro; IPR001137; -  
DR Pfam; PF00457; GLYCO-HYDROL\_11; 1.  
DR PRINTS; PR00911; GLYHDXASE11.  
DR PROSITE; PS00776; GLYCOSYL-HYDROL\_F11\_1; 1.  
DR PROSITE; PS00777; GLYCOSYL-HYDROL\_F11\_2; 1.  
KM Xylan degradation; Hydrolase; Glycosidase.  
FT CHAIN 39 228 ENDO-1,4-BETA-XYLANASE.  
SQ SEQUENCE 228 AA; 24493 MW; EAA15233052E744A CRC64;

Query Match  
Best Local Similarity 67.2%; Score 39; DB 2; Length 228;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTMGOYNAV 9  
DB 169 KTFNOYMSV 177

RESULT 14  
Q9RXG5 PRELIMINARY; PRT; 236 AA.

AC Q9RXG5;  
DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
DE CYTOCHROME C-TYPE BIOGENESIS PROTEIN, HEME EXPORTER PROTEIN C.  
GN DR0348.  
OS Deinococcus radiodurans.  
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.  
OX NCBI\_TaxID=1299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=R1;  
RA MEDLINE=20036896; PubMed=10567266;  
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,

RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,  
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
RA Fraser C.M.;  
RT "Genome sequence of the radioresistant bacterium Deinococcus  
RT radiodurans R1";  
RL Science 286:1571-1577(1999).  
DR EMBL; AE001895; AAF09930.1; -  
DR TIGR; DR0348; -  
DR InterPro; IPR002541; -  
DR InterPro; IPR003557; -  
DR Pfam; PF01578; CytC.asm; 1.  
DR PRINTS; PR01386; CCMCBIOGENSIS.  
SQ SEQUENCE 236 AA; 26591 MW; 18E9BD37E12CE711 CRC64;

Query Match  
Best Local Similarity 83.3%; Score 39; DB 2; Length 236;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TWGOYNAV 7  
DB 101 TWGOYNAV 106

RESULT 15  
P70874 PRELIMINARY; PRT; 259 AA.

AC P70874;  
DT 01-FEB-1997 (TREMblrel. 02, Created)  
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)  
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
DE SPORE CORTEX-LYTIC ENZYME PREPEPTIDE PRECURSOR (GERMINATION-SPECIFIC  
DE AMIDASE) (SCLE).  
GN SLEB.  
OS Bacillus cereus.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1396;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-52; 64-80; 83-94 AND 215-234.  
RC STRAIN=IFO 13597;  
RX MEDLINE=96359394; PubMed=8752358;  
RA Moriyama R., Kudoh S., Miyata S., Nonobe S., Hattori A., Makino S.;  
RT "A germination-specific spore cortex-Lytic enzyme from Bacillus cereus  
RT spores: cloning and sequencing of the gene and molecular  
RT characterization of the enzyme";  
RL J. Bacteriol. 178:5330-5332(1996).  
CC -1- FUNCTION: PROBABLE N-ACETYLTRANSFERASE-L-ALANINE AMIDASE. DEGRADATES  
CC DECOATED SPORES DERIVED FROM VARIOUS ORGANISMS. PRESENT IN DORMANT  
CC SPORES IN ITS ACTIVE FORM.  
CC -1- SUBCELLULAR LOCATION: SPORE-BOUND.  
DR EMBL; D63645; BAA09800.1; -  
DR InterPro; IPR002477; -  
DR Pfam; PF01471; PG\_binding\_1; 1.  
KM Hydrolase; Cell wall; Signal; Repeat.  
FT SIGNAL 1 32  
FT CHAIN 33 259 SPORE-CORTEX-LYTIC ENZYME.  
FT DOMAIN 33 100 MOTIF A.  
FT DOMAIN 65 96 2 X 8 AA APPROXIMATE REPEATS.  
FT REPEAT 65 72 1.  
FT REPEAT 88 96 2.  
FT DOMAIN 154 223 MOTIF B.  
SQ SEQUENCE 259 AA; 28257 MW; 36F266D32EDA544E CRC64;

Query Match  
Best Local Similarity 71.4%; Score 39; DB 2; Length 259;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 WGOYNAV 9  
DB 111

Db 70 WGYWAL 76

RESULT 16

ID 055126 PRELIMINARY; PRT; 281 AA.

AC 055126; 01-JUN-1998 (TREMBlrel. 06, Created)

DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)

DR 01-NOV-1999 (TREMBlrel. 12, Last annotation update)

DE 4-NTTROPHEMYLPHOSPHATASE DOMAIN AND NON-NEURONAL SNAP25-LIKE PROTEIN 2

DE (NIPSNAP2 PROTEIN).

GN GBAS OR NIPSNAP2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Seroussi E., Pan H.Q., Kedra D., Roe B., Dumanaki J.P.;

RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ001261; CAA04635.1; -

DR MGD; MGI:1278343; Gbas.

SQ SEQUENCE 281 AA; 32932 MW; 260D52675BF1CA7E CRC64;

Query Match 67.2%; Score 39; DB 11; Length 281;

Best Local Similarity 83.3%; Pred. No. 68;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WGYWYA 8

Db 197 WGYWYA 202

RESULT 17

ID 043801 PRELIMINARY; PRT; 285 AA.

AC 043801; 01-JUN-1998 (TREMBlrel. 06, Created)

DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)

DR 01-MAY-2000 (TREMBlrel. 13, Last annotation update)

DE NIPSNAP2 PROTEIN.

GN NIPSNAP2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Seroussi E., Pan H.Q., Kedra D., Roe B., Dumanaki J.P.;

RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ001259; CAA04633.1; -

DR InterPro; IPR001066; -

DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; UNKNOWN\_1.

SQ SEQUENCE 285 AA; 33769 MW; B83CD282EA759D2D CRC64;

Query Match 67.2%; Score 39; DB 4; Length 285;

Best Local Similarity 83.3%; Pred. No. 69;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WGYWYA 8

Db 201 WGYWYA 206

RESULT 18

ID 075323 PRELIMINARY; PRT; 286 AA.

AC 075323; 01-JUL-1997 (TREMBlrel. 04, Created)

DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)

DR 01-NOV-1998 (TREMBlrel. 08, Created)

DR 01-NOV-1998 (TREMBlrel. 08, Last sequence update)

DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE GBAS.

GN GBAS.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98277458; PubMed=9615231;

RA Wang X.Y., Smith D.T., Liu W., James C.D.;

RT "GBAS, a novel gene encoding a protein with tyrosine phosphorylation sites and a transmembrane domain, is co-amplified with Esfr.;"

RL Genomics 49:448-451(1998).

DR EMBL; AF029786; AAC29002.1; -

SQ SEQUENCE 286 AA; 33742 MW; 7ED85297E4DC9D08 CRC64;

Query Match 67.2%; Score 39; DB 4; Length 286;

Best Local Similarity 83.3%; Pred. No. 69;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WGYWYA 8

Db 202 WGYWYA 207

RESULT 19

ID 09P058 PRELIMINARY; PRT; 288 AA.

AC 09P058; 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DR 01-MAY-2001 (TREMBlrel. 16, Last annotation update)

DE NIPSNAP2 PROTEIN (FRAGMENT).

GN NIPSNAP2.

OS Brachydanio rerio (Zebrafish) (Zebra danio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes; Cyprinidae; Rasbora; Danio.

OX NCBI\_TaxID=7955;

RN [1]

RP SEQUENCE FROM N.A.

RA Kedra D., Dumanaki J.P.;

RT "Cloning of NIPSNAP gene orthologues in Danio rerio and Drosophila melanogaster."

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ249797; CAB56702.1; -

DR InterPro; IPR001781; -

DR PRODOM; PD000094; -; 1.

FT NON\_TER 1 1

SQ SEQUENCE 288 AA; 33667 MW; 66EDDBB045C6288C CRC64;

Query Match 67.2%; Score 39; DB 13; Length 288;

Best Local Similarity 83.3%; Pred. No. 69;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WGYWYA 8

Db 204 WGYWYA 209

RESULT 20

ID 008346 PRELIMINARY; PRT; 335 AA.

AC 008346; 01-JUL-1997 (TREMBlrel. 04, Created)

DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)

DR 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE XYLANASE PRECURSOR (EC 3.2.1.8) (ENDO-1,4-BETA-XYLANASE) (1,4-BETA-D-

DE XYLAN XYLANOXYDROLASE).

GN XYLIT.



OS Streptomyces thermoviolaceus.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1952;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OPC-520;  
 RX MEDLINE=97176398; PubMed=9023944;  
 RA Tsujibo H., Ohtsuki T., Ito T., Yamazaki I., Miyamoto K., Sugiyama M.,  
 Yamori Y.;  
 RT "Cloning and sequence analysis of genes encoding xylanases and acetyl  
 xylan esterase from Streptomyces thermoviolaceus OPC-520."  
 RL Appl. Environ. Microbiol. 63:661-664(1997).  
 CC -1- CATALYTIC ACTIVITY: ENDOMYDOLYSIS OF 1,4-BETA-D-XYLOSIDIC  
 LINKAGES IN XYLANS.  
 DR EMBL: D85897; BAI19778.1; -  
 DR HSSP: P09850; 1XB.  
 DR InterPro: IPR001137; -  
 DR Pfam: PF00457; Glyco\_hydro\_11; 1.  
 DR PRINTS: PR00911; GLYCOSYL\_HYDROL\_F11\_1;  
 DR PROSITE: PS00776; GLYCOSYL\_HYDROL\_F11\_1; 1.  
 DR PROSITE: PS00777; GLYCOSYL\_HYDROL\_F11\_2; 1.  
 KW Signal: xylan degradation; Hydrolase; Glycosidase.  
 FT SIGNAL 1 29  
 FT CHAIN 30 335  
 FT SEQUENCE 335 AA; 35280 MW; B9C9A0C2BA02341 CRC64;

Query Match 67.2%; Score 39; DB 2; Length 335;  
 Best Local Similarity 66.7%; Pred. No. 81;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 KTMGOYAV 9  
 Db 170 KTFNOYMSV 178

RESULT 21  
 O9RMM4  
 ID O9RMM4 PRELIMINARY; PRT; 335 AA.  
 AC O9RMM4;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE XYLANASE B PRECURSOR.  
 CN XYNB.  
 OS Streptomyces thermocyanoviolaceus.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=106355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KCM 40049;  
 RA Shin J.-H., Choi J.-H., Lee O.-S., Joo G.-J., Rhee I.-K.;  
 RT "Streptomyces thermocyanoviolaceus xynb (xylanase B)."  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF194025; AF04601.1; -  
 DR HSSP: P09850; 1XB.  
 DR InterPro: IPR001137; -  
 DR Pfam: PF00457; Glyco\_hydro\_11; 1.  
 DR PRINTS: PR00911; GLYDRLASE11.  
 DR PROSITE: PS00776; GLYCOSYL\_HYDROL\_F11\_1; 1.  
 DR PROSITE: PS00777; GLYCOSYL\_HYDROL\_F11\_2; 1.  
 KW Signal: xylan degradation; Hydrolase; Glycosidase.  
 FT SIGNAL 1 40  
 FT CHAIN 41 335  
 FT SEQUENCE 335 AA; 35328 MW; D7DED7BAFCB52710 CRC64;

Query Match 67.2%; Score 39; DB 2; Length 335;  
 Best Local Similarity 66.7%; Pred. No. 81;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KTMGOYAV 9  
 Db 170 KTFNOYMSV 178

RESULT 22  
 O9W616  
 ID O9W616 PRELIMINARY; PRT; 358 AA.  
 AC O9W616;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE SIGNALING MOLECULE LEFTY1.  
 CN LEFTY1.  
 OS Brachydanio rerio (Zebrafish) (Zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Osteichthys;  
 OC Cypriniformes; Cyprinidae; Rasbora; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bisgrove B.W., Essner J.J., Yost H.J.;  
 RT "Regulation of axis development by antagonism of lefty and nodal  
 signaling in the midline."  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO TRANSFORMING GROWTH FACTOR (TGF) BETA FAMILY.  
 DR EMBL: AF132444; A043388.1; -  
 DR HSSP: P01137; 1KLA.  
 DR InterPro: IPR001111; -  
 DR InterPro: IPR001839; -  
 DR Pfam: PF00019; TGF-beta; 1.  
 DR Pfam: PF00688; TGF-beta; 1.  
 DR PROSITE: PD00357; -1.  
 DR PROSITE: PS00250; TGF-BETA; 1.  
 DR SMART: SM00204; TGF-beta; 1.  
 KW Glycoprotein.  
 SQ SEQUENCE 358 AA; 41018 MW; F7A86FE828092A8 CRC64;

Query Match 67.2%; Score 39; DB 13; Length 358;  
 Best Local Similarity 62.5%; Pred. No. 86;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Oy 2 TWGOYAV 9  
 Db 278 TWOTWIT 285

RESULT 23  
 O9PW55  
 ID O9PW55 PRELIMINARY; PRT; 358 AA.  
 AC O9PW55;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE ANTRIN.  
 OS Brachydanio rerio (Zebrafish) (Zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Osteichthys;  
 OC Cypriniformes; Cyprinidae; Rasbora; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=99065494; PubMed=9847237;  
 RA Thise C., Thise B.;  
 RT "Antivin, a novel and divergent member of the TGF-beta superfamily,  
 RT negatively regulates mesoderm induction."  
 RL Development 126:229-240(1999).  
 CC -1- SIMILARITY: TO TRANSFORMING GROWTH FACTOR (TGF) BETA FAMILY.  
 DR EMBL: AF172069; A047114.1; -  
 DR HSSP: P01137; 1KLA.  
 DR InterPro: IPR001111; -  
 DR InterPro: IPR001839; -

DR Pfam: PF00019; TGF-beta; 1.  
 DR Pfam: PF00688; TGF-propeptide; 1.  
 DR PROSITE: PS00250; TGF\_BETA; 1.  
 DR SMART: SM00204; TGFb; 1.  
 KW Glycoprotein.  
 SQ SEQUENCE 358 AA; 41041 MW; 74D3036F18D2458F CRC64;

Query Match 67.2%; Score 39; DB 13; Length 358;  
 Best Local Similarity 62.5%; Pred. No. 86;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TWGOYVAV 9  
 |||||  
 Db 278 TWTOYWI 285

RESULT 24  
 Q9W617 PRELIMINARY; PRT; 362 AA.  
 AC Q9W617;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE SIGNALING MOLECULE LEFTY2.  
 GN LEFTY.  
 OS Brachydanio rerio (zebrafish) (Zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 OC Cypriniformes; Cyprinidae; Rasbora; Danio.  
 OC NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Biströme B.W., Essner J.J., Yost H.J.;  
 RT "Regulation of axis development by antagonism of Lefty and nodal  
 RT signaling in the midline."  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- Similarity: TO TRANSFORMING GROWTH FACTOR (TGF) BETA FAMILY.  
 DR EMBL; AF134445; AAD34389.1; -  
 DR HSSP; P08112; 2TGI.  
 DR InterPro: IPR001111; -  
 DR InterPro: IPR001839; -  
 DR Pfam: PF00019; TGF-beta; 1.  
 DR Pfam: PF00688; TGF-propeptide; 1.  
 DR PROSITE: PS00250; TGF\_BETA; 1.  
 DR SMART; SM00204; TGFb; 1.  
 KW Glycoprotein.  
 SQ SEQUENCE 362 AA; 41459 MW; 0C12B8CB3CEB6007 CRC64;

Query Match 67.2%; Score 39; DB 13; Length 362;  
 Best Local Similarity 62.5%; Pred. No. 87;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TWGOYVAV 9  
 |||||  
 Db 281 TWTOYWI 288

RESULT 25  
 Q9DFC6 PRELIMINARY; PRT; 366 AA.  
 AC Q9DFC6;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE TGF-BETA FAMILY MEMBER LEFTY-A.  
 GN Xenopus laevis (African clawed frog).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=20341055; PubMed=10882517;  
 RA Branford W.W., Essner J.J., Yost H.J.;  
 RT "Regulation of gut and heart left-right asymmetry by context-dependent  
 RT interactions between Xenopus Lefty and BMP4 signaling."  
 RL Dev. Biol. 223:291-306(2000).  
 DR EMBL; AF283562; AAG10034.1; -  
 SQ SEQUENCE 366 AA; 41434 MW; E65CEC306F4B5ED5 CRC64;

Query Match 67.2%; Score 39; DB 13; Length 366;  
 Best Local Similarity 62.5%; Pred. No. 88;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TWGOYVAV 9  
 |||||  
 Db 284 TWTOYWI 291

RESULT 26  
 Q9DFC5 PRELIMINARY; PRT; 367 AA.  
 AC Q9DFC5;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE TGF-BETA FAMILY MEMBER LEFTY-B.  
 GN Xenopus laevis (African clawed frog).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Branford W.W., Essner J.J., Yost H.J.;  
 RT "Regulation of gut and heart left-right asymmetry by context-dependent  
 RT interactions between Xenopus Lefty and BMP4 signaling."  
 RL Dev. Biol. 223:291-306(2000).  
 DR EMBL; AF283563; AAG10035.1; -  
 SQ SEQUENCE 367 AA; 41604 MW; DBA7E85752B7FB3A CRC64;

Query Match 67.2%; Score 39; DB 13; Length 367;  
 Best Local Similarity 62.5%; Pred. No. 88;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TWGOYVAV 9  
 |||||  
 Db 284 TWTOYWI 291

RESULT 27  
 Q9DD36 PRELIMINARY; PRT; 367 AA.  
 AC Q9DD36;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE XANTIVIN (LEFTY-RELATED FACTOR XATV).  
 GN XANTIVIN.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tanegashima K., Yokota C., Takahashi S., Asashima M.;  
 RT "Expression cloning of Xantivin, a Xenopus lefty/xantivin-related gene,  
 RT involved in the regulation of activin signaling during mesoderm  
 induction."

```

RL Mech. Dev. 99:3-14(2000).
RN
RP SEQUENCE FROM N.A.
RA Cheng A.M.S., Thisse B., Thisse C., Wright C.V.E.;
RT "The lefty-related factor Xatv acts as a feedback inhibitor of nodal
RL signaling in mesoderm induction and L-R axis development in Xenopus.";
RL Development 0:0-0(2000).
DR EMBL: AB038010; BAB12725.1;
DR EMBL: AF209744; AAG35771.1;
SQ SEQUENCE 367 AA; 41504 MW; DEAA90275BC8574A CRC64;

Query Match
Best Local Similarity 62.5%; Score 39; DB 13; Length 367;
Pred. No. 88;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TWGQYNAV 9
DB 284 TWGQYNAV 291

RESULT 28
ID 050200 PRELIMINARY; PRT; 469 AA.
AC 050200;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE AMYLASE (EC 3.2.1.1) (ALPHA-AMYLASE) (1,4-ALPHA-D-GLUCAN
DE GLUCANOHYDROLASE) (GLYCOGENASE).
GN AMY.
OS Thermococcus sp. R3.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.
OX NCBI_TaxID=55421;
RN [1]
RP SEQUENCE FROM N.A.
RA Jones R.A., Patel B., Beascham I.R.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -1 CATALYTIC ACTIVITY: ENDOTHERMOLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
DR EMBL: AF017454; AAB87860.1;
DR HSSP: P06278; IVJS.
DR InterPro: IPR000461;
DR Pfam: PF00128; alpha-amylase; 1.
KW Hydrolase; Glycosidase.
SQ SEQUENCE 469 AA; 52756 MW; C3B0614FBA4DB3E2 CRC64;

Query Match
Best Local Similarity 71.4%; Score 39; DB 1; Length 469;
Pred. No. 11e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYNAV 7
DB 204 KSMQYNAV 210

RESULT 29
ID 092PS2 PRELIMINARY; PRT; 545 AA.
AC 092PS2;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE PUTATIVE OLIGOPEPTIDE TRANSPORT PROTEIN.
GN F14H20.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Rongling C.M., Benito M.,
RA Carrera A.J., Greasy T.H., Buell C.R., Town C.D., Mierman W.C.,
RA Fraser C.M., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC F14H20 genomic sequence.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC006532; AAD20094.1;
DR InterPro: IPR000109;
DR Pfam: PF00854; PTR2; 2.
DR PROSITE: PS01022; PTR2.1; 1.
DR PROSITE: PS01023; PTR2.2; UNKNOWN.1.
SQ SEQUENCE 545 AA; 60879 MW; E69740E3A53FEA5F CRC64;

Query Match
Best Local Similarity 57.1%; Score 39; DB 10; Length 545;
Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 WGYNAV 9
DB 111 WGYNAV 117

RESULT 30
ID 09SDN4 PRELIMINARY; PRT; 559 AA.
AC 09SDN4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE AMINO ACID/PEPTIDE TRANSPORTER.
GN PTR2.
OS Prunus dulcis (Almond) (Prunus amygdalus).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Rosales; Rosaceae; Prunus.
OX NCBI_TaxID=3755;
RN [1]
RP SEQUENCE FROM N.A.
RA Campalans A., Pages M., Messegue R.;
RT "Identification of differentially expressed genes during dehydration
RT in almond (Prunus amygdalus) using the cDNA-AFLP technique.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF213936; AAR20002.1;
DR InterPro: IPR000109;
DR Pfam: PF00854; PTR2; 1.
DR PROSITE: PS01022; PTR2.1; 1.
DR PROSITE: PS01023; PTR2.2; 1.
SQ SEQUENCE 559 AA; 61604 MW; AAB9D08BD085C9A3 CRC64;

Query Match
Best Local Similarity 57.1%; Score 39; DB 10; Length 559;
Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 WGYNAV 9
DB 110 WGYNAV 116

Search completed: July 3, 2001, 10:28:10
Job time: 330 sec

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: July 3, 2001, 10:00:55 ; Search time 23.85 Seconds  
(without alignments)  
28.745 Million cell updates/sec

Title: US-09-214-836-1  
Perfect score: 58  
Sequence: 1 KTWGQYNAV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pirl:\*  
2: pirl:\*  
3: pirl:\*  
4: pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	91.4	662	2 I38400	melanoma-associated
2	53	91.4	668	2 A41234	melanocyte-specific
3	49	84.5	626	2 S53871	Pmel 17 protein -
4	42	72.4	305	2 C69708	spore cortex-lytic
5	42	72.4	315	2 T07314	cytochrome c-type
6	40	69.0	264	2 H85861	hypothetical prote
7	40	69.0	276	2 A72451	probable lactose t
8	40	69.0	281	2 A82104	conserved hypothet
9	40	69.0	303	1 S52775	hypothetical prote
10	40	69.0	400	2 S76446	hypothetical prote
11	40	69.0	549	2 H64992	hypothetical prote
12	40	69.0	807	2 F64844	ydds protein precu
13	40	69.0	807	2 F85648	probable outer mem
14	39	67.2	236	2 A75350	tycoxanthin chloro
15	39	67.2	236	2 A75350	hypothetical prote
16	39	67.2	256	2 B70750	protein-tyrosine-p
17	39	67.2	335	1 A39862	probable peptide/a
18	39	67.2	545	2 A84432	hypothetical prote
19	39	67.2	568	2 E96648	histidine transpor
20	39	67.2	585	2 C84432	histidine transpor
21	39	67.2	586	2 S46236	probable membrane
22	39	67.2	1081	2 B81303	hypothetical prote
23	38	65.5	209	2 S75029	sulfate/rhiosulfat
24	38	65.5	277	1 ORECS7	hypothetical prote
25	38	65.5	277	2 B85885	hypothetical prote
26	38	65.5	290	2 C86097	4-hydroxybenzoate-
27	38	65.5	290	2 JC2316	hypothetical prote
28	38	65.5	569	2 T22928	hypothetical prote
29	38	65.5	622	2 H64447	hypothetical prote

30	38	65.5	936	2 B64567	cytochrome c bioge
31	38	65.5	936	2 H71862	probable cytochrom
32	37	63.8	240	1 JS0591	endo-1,4-beta-xyla
33	37	63.8	241	2 T37005	endo-1,4-beta-xyla
34	37	63.8	293	2 T11969	cytochrome c-type
35	37	63.8	301	2 D82040	cysQ protein VC272
36	37	63.8	306	1 S25309	cytochrome c-type
37	37	63.8	312	2 S78242	cytochrome c-type
38	37	63.8	319	1 S73290	cytochrome c-type
39	37	63.8	320	2 T07572	hypothetical prote
40	37	63.8	322	2 T06955	probable cytochrom
41	37	63.8	324	2 D70545	probable cytochrom
42	37	63.8	325	2 F65088	hypothetical prote
43	37	63.8	327	1 S72913	cytochrome c-type
44	37	63.8	334	1 S74957	cytochrome c-type
45	37	63.8	353	2 T07998	cytochrome c-type

## ALIGNMENTS

RESULT 1  
I38400 melanoma-associated ME20 antigen (me20m) - human  
N:Alternate names: melanoma antigen 25  
C:Species: Homo sapiens (man)  
C>Date: 01-Nov-1996 #sequence\_revision 01-Nov-1996 #text\_change 01-Dec-2000  
C:Accession: I38400; A53668; A55753  
R:Marsh, G.A.; Marken, J.S.; Neubauer, M.; Aruffo, A.; Hellstrom, I.; Hellstrom, K.; DNA Cell Biol. 13, 87-95, 1994  
A>Title: Cloning and expression of the gene for the Melanoma-Associated ME20 Antigen.  
A:Reference number: I38400; MUID:94235165  
A:Accession: I38400  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-662 <RES>  
A:Cross-references: EMBL:U01874; NID:9494939; PID:AAA18479.1; PID:9494940  
R:Adema, G.J.; de Boer, A.J.; Vogel, A.M.; Loenen, W.A.M.; Figdor, C.G. J. Biol. Chem. 269, 20126-20133, 1994  
A>Title: Molecular characterization of the melanocyte lineage-specific antigen gp100.  
A:Reference number: A53668; MUID:94327568  
A:Accession: A53668  
A:Molecule type: mRNA  
A:Residues: 1-592,594-662 <ADE>  
A:Cross-references: GB:S73003; NID:9639589; PID:AMC60634.1; PID:9639590  
R:Kwakami, Y.; Eliyahu, S.; Delgado, C.H.; Robins, P.F.; Sakaguchi, K.; Appella, E. Proc. Natl. Acad. Sci. U.S.A. 91, 6458-6462, 1994  
A>Title: Identification of a human melanoma antigen recognized by tumor-infiltrating  
A:Reference number: A55753; MUID:94294401  
A:Accession: A55753  
A>Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-161,163-592,594-662 <KAW>  
C:Keywords: glycoprotein

Query Match 91.4%; Score 53; DB 2; Length 662;  
Best Local Similarity 88.9%; Pred. No. 0.88;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KTWGQYNAV 9  
Db 154 KTWGQYNAV 162

RESULT 2  
A41234 melanocyte-specific protein Pmel-17 precursor - human  
C:Species: Homo sapiens (man)  
C>Date: 19-Jun-1992 #sequence\_revision 19-Jun-1992 #text\_change 30-Sep-1993  
C:Accession: A41234  
R:Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.; Gilbert, D.J.; Jenkins, Proc. Natl. Acad. Sci. U.S.A. 88, 9228-9232, 1991

A:Title: A melanocyte-specific gene, Pmel 17, maps near the silver coat color locus on H  
 A:Reference number: A41234; MUID:92021023  
 A:Accession: A41234  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-668 <KMO>  
 A:Cross-references: GB:M77348

Query Match 91.4%; Score 53; DB 2; Length 668;  
 Best Local Similarity 88.9%; Pred. No. 0.89;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYMAV 9  
 |||||  
 DB 154 KTWGQYMAV 162

RESULT 3  
 S53871  
 Pmel 17 protein - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 05-Nov-1999  
 C:Accession: S53871  
 R:Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.; Chintamani, C.; Bennett, D.; Pick  
 Nucleic Acids Res. 23, 154-158, 1995  
 A:Title: Mouse silver mutation is caused by a single base insertion in the putative cyto  
 A:Reference number: S53871; MUID:95175558  
 A:Accession: S53871  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-626 <KMO>  
 A:Cross-references: GB:U14133; MID:9887940; PIDN:AAA69538.1; PID:9887941

Query Match 84.5%; Score 49; DB 2; Length 626;  
 Best Local Similarity 77.8%; Pred. No. 3.4;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYMAV 9  
 |||||  
 DB 154 KTWGQYMAV 162

RESULT 4  
 C69708  
 spore cortex-lytic enzyme prepeptide [imported] - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 21-Jul-2000  
 C:Accession: C69708; T44770  
 R:Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Ch  
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
 Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portelle  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon  
 A:Authors: Scheich, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
 tkeuchl, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunori, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033  
 A:Accession: C69708  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-305 <KUN>  
 A:Cross-references: GB:Z99115; GB:Z99116; GB:AL009126; MID:g2634723; PIDN:CBM14225.1; PI  
 A:Experimental source: strain 168  
 R:Moriyama, R.; Hattori, A.; Miyata, S.; Kudoh, S.; Makino, S.

J. Bacteriol. 178, 6059-6063, 1996  
 A:Title: A gene (slrB) encoding a spore cortex-lytic enzyme from Bacillus subtilis an  
 A:Reference number: Z22836; MUID:96427343  
 A:Accession: T44770  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-305 <MOR>  
 A:Cross-references: EMBL:D79978; MID:g1688021; PIDN:BAAL1473.1; PID:g1688023  
 A:Experimental source: strain 168  
 A:Genetics: slrB

Query Match 72.4%; Score 42; DB 2; Length 305;  
 Best Local Similarity 85.7%; Pred. No. 20;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WGYWMAV 9  
 |||||  
 DB 67 WGYWMAV 73

RESULT 5  
 T07314  
 cytochrome c-type synthesis protein homolog - Chlorella vulgaris chloroplast  
 C:Species: Chlorella vulgaris  
 C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000  
 C:Accession: T07314  
 R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Na  
 Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997  
 A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga C  
 A:Reference number: T15985; MUID:97303241  
 A:Accession: T07314  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-315 <MAK>  
 A:Cross-references: EMBL:AB001684; MID:g2224352; PIDN:BAAS7962.1; PID:g2224478  
 A:Genetics: A:Gene: ycf5  
 A:Genome: chloroplast  
 C:Superfamily: cytochrome c-type synthesis protein  
 C:Keywords: chloroplast

Query Match 72.4%; Score 42; DB 2; Length 315;  
 Best Local Similarity 62.5%; Pred. No. 20;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYMA 8  
 |||||  
 DB 243 ETWGNWMS 250

RESULT 6  
 H85861  
 hypothetical protein Z3480 [imported] - Escherichia coli (strain O157:H7)  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
 C:Accession: H85861  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
 iller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamousis, K.; Apoda  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: AB5480; MUID:21074935; PMID:11206551  
 A:Accession: H85861  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-264 <STO>  
 A:Cross-references: GB:AC005174; MID:g12516559; PIDN:AAG57356.1; GSPDB:GN00145; UWGP:  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics: A:Gene: Z3480

Query Match 69.0%; Score 40; DB 2; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGOYW 7  
 |||||  
 Db 125 WGOYW 129

## RESULT 7

A72451  
 Probable lactose transport system permease protein APE2253 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: A72451  
 R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum A:Reference number: A72450; MUID:99310339  
 A:Accession: A72451  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-276 <KAN>  
 A:Cross-references: DDBJ:AF000064; NID:95105945; PIDN:BA81265.1; PID:95105954  
 A:Experimental source: strain K1  
 C:Genetics:  
 A:Gene: APE2253  
 C:Superfamily: Inner membrane protein 4gpa

Query Match 69.0%; Score 40; DB 2; Length 276;  
 Best Local Similarity 55.6%; Pred. No. 36;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYMW 9  
 :|||:|  
 Db 215 RTWGQWML 223

## RESULT 8

A82104  
 conserved hypothetical protein VC2229 [Imported] - Vibrio cholerae (strain N16961 serogr C:Species: Vibrio cholerae  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C:Accession: A82104  
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A:Reference number: A82035; MUID:20406833  
 A:Accession: A82104  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-281 <HEI>  
 A:Cross-references: GB:AE004294; GB:AE003852; NID:99656774; PIDN:AAF95373.1; GSPDB:GN001 C:Genetics:  
 A:Gene: VC2229  
 A:Map position: 1  
 C:Superfamily: hypothetical protein H11037

Query Match 69.0%; Score 40; DB 2; Length 281;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGOYW 7  
 |||||  
 Db 99 WGOYW 103

RESULT 9  
 S52775  
 hypothetical protein 2 - Chloroflexus aurantiacus  
 C:Species: Chloroflexus aurantiacus  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
 C:Accession: S52775  
 R:Niedermeyer, G.; Shiozawa, J.A.; Lottspeich, F.; Felck, R.G.  
 FEBS Lett. 342, 61-65, 1994  
 A:Title: The primary structure of two chlorosome proteins from Chloroflexus aurantiac A:Reference number: S43678; MUID:94192803  
 A:Accession: S52775  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-305 <NEI>  
 A:Cross-references: EMBL:Z34000; NID:9496485; PIDN:CAA83969.1; PID:9496488  
 A:Note: only a part of the coding sequence is given in this paper  
 A:Note: the nucleotide sequence was submitted to the EMBL data library, May 1994  
 C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0279

Query Match 69.0%; Score 40; DB 1; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGOYW 7  
 |||||  
 Db 246 WGOYW 250

## RESULT 10

S76446  
 hypothetical protein - Synechocystis sp. (strain PCC 6803)  
 C:Species: Synechocystis sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
 C:Accession: S76446  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O, K.; Okumura, S.; Shingo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis.  
 A:Reference number: S74322; MUID:97061201  
 A:Accession: S76446  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-400 <KAN>  
 A:Cross-references: EMBL:D90915; GB:AB001339; NID:91653604; PIDN:BA18575.1; PID:d101 A:Note: the nucleotide sequence was submitted to the EMBL data library, June 1996

Query Match 69.0%; Score 40; DB 2; Length 400;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGOYW 7  
 |||||  
 Db 335 WGOYW 339

## RESULT 11

H64992  
 hypothetical protein b2226 - Escherichia coli (strain K-12)  
 C:Species: Escherichia coli  
 C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 04-Mar-2000  
 C:Accession: H64992  
 R:Battner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617  
 A:Accession: H64992  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
A:Residues: 1-549 <BLAT>  
A:Cross-References: GB:AE000312; GB:U00096; NID:g178855; PIDN:AC75286.1; PID:g1788557;  
A:Experimental source: strain K-12, substrain MG1655  
C:Superfamily: Escherichia coli hypothetical protein b2226

Query Match 69.0%; Score 40; DB 2; Length 549;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOYW 7  
|||  
DB 125 WGOYW 129

RESULT 12  
F64844  
Ycds protein precursor - Escherichia coli  
C:Species: Escherichia coli  
C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 04-Mar-2000  
C:Accession: F64844  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: F64844  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-807 <BLAT>  
A:Cross-References: GB:AE000204; GB:U00096; NID:g1787256; PIDN:AC74109.1; PID:g1787261;  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: ycds  
C:Superfamily: Escherichia coli ycds protein  
F:1-76/Domain: signal sequence #status predicted <SIG>  
F:27-807/Product: ycds protein #status predicted <MAT>

Query Match 69.0%; Score 40; DB 2; Length 807;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOYW 7  
|||  
DB 314 WGOYW 318

RESULT 13  
F85647  
Probable outer membrane protein ycds [imported] - Escherichia coli (strain O157:H7)  
C:Species: Escherichia coli  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
C:Accession: F85647  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grothbeck, E.J.; Davis, N.W.; Lm, A.; Dinalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: F85647  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-807 <STO>  
A:Cross-References: GB:AE00174; NID:g12514389; PIDN:AC655642.1; GSPDB:GN00145; UWGP:Z15  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: ycds  
C:Superfamily: Escherichia coli ycds protein

Query Match 69.0%; Score 40; DB 2; Length 807;  
Best Local Similarity 100.0%; Pred. No. 97;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 3 WGOYW 7  
|||  
DB 314 WGOYW 318

RESULT 14  
S46301  
fucoxanthin chlorophyll a/c-binding light-harvesting protein - Isochrysis galbana  
N:Alternate names: FCP protein  
C:Species: Isochrysis galbana  
C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 05-Nov-1999  
C:Accession: S46301  
R:Larocque, J.; Henry, D.; Wyman, K.; Suenik, A.; Falkowski, P.  
Plant Mol. Biol. 25, 355-368, 1994  
A:Title: Cloning and nucleotide sequence of a cDNA encoding a major fucoxanthin-, chl  
e family.  
A:Reference number: S46301; MUID:94325461  
A:Accession: S46301  
A:Molecule type: mRNA  
A:Residues: 1-208 <LARA>  
A:Cross-References: EMBL:X77333; NID:g535080; PIDN:CA54547.1; PID:g535081  
A:Experimental source: cultivar DUN  
C:Keywords: light-harvesting complex

Query Match 67.2%; Score 39; DB 2; Length 208;  
Best Local Similarity 83.3%; Pred. No. 40;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 TWGOYW 7  
|||  
DB 198 TWGOYW 203

RESULT 15  
A75330  
cytochrome c-type biogenesis protein, heme exporter protein C - Deinococcus radiodurans  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: A75330  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
M.; Shen, M.; Vamathevan, J.U.; Lam, P.; McDonald, L.; Uitterlind, T.; Zalewski, C.;  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: A75330  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-236 <WHI>  
A:Cross-References: GB:AE001895; GB:AE000513; NID:g6458024; PIDN:AAF09930.1; PID:g645  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0348  
A:Map position: 1  
C:Superfamily: helC protein

Query Match 67.2%; Score 39; DB 2; Length 236;  
Best Local Similarity 83.3%; Pred. No. 45;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 TWGOYW 7  
|||  
DB 101 TWGOYW 106

RESULT 16  
B70750  
hypothetical protein RV0090 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis



C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 28-Jul-2000  
 C/Accession: B70750  
 R/Cole, S.T.; Brosch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
 A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; M0ID:98295987  
 A/Accession: B70750  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-256 <COL>  
 A/Cross-references: GB:274410; GB:AL123456; NID:g3261600; PIDN:CAA98926.1; PID:e249404;  
 C:Experimental source: strain H37Rv  
 C:Genetics:  
 A:Gene: RV0090  
 C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV0090

Query Match 67.2%; Score 39; DB 2; Length 256;  
 Best Local Similarity 66.7%; Pred. No. 48;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTMGOYAV 9  
 |||:|:  
 Db 177 KRMGEYFV 185

RESULT 17  
 A39862  
 protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 1 - yeast (Saccharomyces cerevisiae)  
 N/Alternate names: protein P0815; protein YDL230w  
 C/Species: Saccharomyces cerevisiae  
 C/Date: 30-Dec-1991 #sequence\_revision 08-Mar-1996 #text\_change 21-Jul-2000  
 C/Accession: A39862; S67793  
 R/Guan, K.; Deschenes, R.J.; Qiu, H.; Dixon, J.E.  
 J. Biol. Chem. 266, 12964-12970, 1991  
 A:Title: Cloning and expression of a yeast protein tyrosine phosphatase.  
 A:Reference number: A39862; M0ID:91302312  
 A/Accession: A39862  
 A:Molecule type: DNA  
 A:Residues: 1-335 <GUA>  
 A/Cross-references: GB:M64062; NID:g172295; PIDN:AAA34923.1; PID:g172296  
 R/Rasmussen, S.W.  
 Submitted to the Protein Sequence Database, July 1996  
 A:Reference number: S67778  
 A/Accession: S67793  
 A:Molecule type: DNA  
 A:Residues: 1-335 <RAS>  
 A/Cross-references: EMBL:274278; NID:g1431387; PIDN:CAA98809.1; PID:g1431388; GSPDB:GN00  
 C:Experimental source: strain S288C  
 C:Genetics:  
 A:Gene: SGD:PTP1; MIPS:YDL230w  
 A/Cross-references: SGD:S0002389; MIPS:YDL230w  
 A:Map position: 4L  
 C:Superfamily: Saccharomyces protein-tyrosine-phosphatase, nonreceptor type 1; protein-t  
 C:Keywords: phosphoprotein; phosphoric monoster hydrolase; tyrosine-specific phosphatase  
 F:52-303/Domain: protein-tyrosine-phosphatase homology <RTP>  
 F:252/Active site: Cys (phosphocysteine intermediate) #status predicted  
 F:258/Binding site: substrate phosphate (Arg) #status predicted

Query Match 67.2%; Score 39; DB 1; Length 335;  
 Best Local Similarity 71.4%; Pred. No. 61;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTMGOYAV 7  
 |||:|:  
 Db 108 KTMGOYFV 114

RESULT 18

A84432  
 probable peptide/amino acid transporter [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C/Accession: A84432  
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,  
 eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; M0ID:20083487  
 A/Accession: A84432  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-545 <STO>  
 A/Cross-references: GB:AE002093; NID:g4406784; PIDN:AAD20094.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g02020  
 A:Map position: 2

Query Match 67.2%; Score 39; DB 2; Length 545;  
 Best Local Similarity 57.1%; Pred. No. 96;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 MGQYMAV 9  
 |||:|:  
 Db 111 WGRYWTI 117

RESULT 19  
 E96648  
 hypothetical protein F19K23.13 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C/Accession: E96648  
 R/Rieolagis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
 ansen, N.F.; Hughes, B.; Hultzer, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marzia  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; M0ID:21016719  
 A/Accession: E96648  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-568 <STO>  
 A/Cross-references: GB:AE005173; NID:g2160144; PIDN:AAB60766.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: F19K23.13  
 A:Map position: 1

Query Match 67.2%; Score 39; DB 2; Length 568;  
 Best Local Similarity 57.1%; Pred. No. 99;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 MGQYMAV 9  
 |||:|:  
 Db 106 WGRYWTI 112

RESULT 20  
 C84432  
 histidine transport protein (PTR2-B) [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C/Accession: C84432  
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.: Koo, H.; Mofiat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: A84420; MUID:20083487  
A:Accession: C84432  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-585 <STO>  
A:Cross-references: GB:AE002093; NID:g4406786; PIDN:AAD20096.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g02040  
A:Map position: 2

Query Match 67.2%; Score 39; DB 2; Length 585;  
Best Local Similarity 57.1%; Pred. No. 1e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 MGQYNAV 9  
|||:  
Db 110 MGRYWT 116

RESULT 21  
S46236  
histidine transport protein - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 21-Jul-2000  
R:Frommer, W.B.; Hummel, S.; Rentsch, D.  
FEBS Lett. 347, 185-189, 1994  
A:Title: Cloning of an *Arabidopsis* histidine transporting protein related to nitrate and  
A:Reference number: S46236; MUID:94307379  
A:Accession: S46236  
A:Molecule type: mRNA  
A:Residues: 1-586 <PRO>  
A:Cross-references: EMBL:X77503; NID:9510237; PIDN:CAA54634.1; PID:9510238  
C:Genetics:  
A:Gene: NTR1  
C:Keywords: amino acid transport; histidine transport; transmembrane protein

Query Match 67.2%; Score 39; DB 2; Length 586;  
Best Local Similarity 57.1%; Pred. No. 1e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 MGQYNAV 9  
|||:  
Db 110 MGRYWT 116

RESULT 22  
B81303  
Probable membrane protein Cj1013c [imported] - *Campylobacter jejuni* (strain NCTC 11168)  
C:Species: *Campylobacter jejuni*  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 31-Mar-2000  
R:Partholl, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
C.W.; O'Neill, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre  
Nature 403, 665-666, 2000  
A:Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hyp  
A:Reference number: A81250; MUID:20150912  
A:Accession: B81303  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1081 <PAR>  
A:Cross-references: GB:AL139077; GB:AL111168; NID:95968444; PIDN:CAB73269.1; PID:9596844  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: Cj1013c

Query Match 67.2%; Score 39; DB 2; Length 1081;  
Best Local Similarity 50.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGQYMA 8  
:|||:  
Db 977 ESMGRYWS 984

RESULT 23  
S75029  
hypothetical protein slr2003 - *Synechocystis* sp. (strain PCC 6803)  
C:Species: *Synechocystis* sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima,  
O.K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocys*  
s.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S75029  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-209 <KAN>  
A:Cross-references: EMBL:D90910; GB:AB001339; NID:g1652956; PIDN:BAA17891.1; PID:g165  
C:Superfamily: *Synechocystis* hypothetical protein slr2003

Query Match 65.5%; Score 38; DB 2; Length 209;  
Best Local Similarity 71.4%; Pred. No. 56;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYMA 7  
|||:  
Db 75 KTWQRYW 81

RESULT 24  
QRECSF  
sulfate/thiosulfate transport protein cyst - *Escherichia coli*  
N:Alternate names: sulfate transport system permease protein cyst  
C:Species: *Escherichia coli*  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Jul-1999  
C:Accession: A35402; B35403  
R:Sirko, A.; Hryniewicz, M.; Hulanicka, D.; Boeck, A.  
J. Bacteriol. 172, 3351-3357, 1990  
A:Title: Sulfate and thiosulfate transport in *Escherichia coli* K-12: nucleotide sequ  
A:Reference number: A35402; MUID:90264334  
A:Accession: A35402  
A:Molecule type: DNA  
A:Residues: 1-277 <SIR>  
A:Cross-references: GB:M32101; GB:M38050; NID:g145657; PIDN:AAA2637.1; PID:g145659  
A:Experimental source: strain K12  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.; Rose, D.J.; Mau, B.; Shaoh, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of *Escherichia coli* K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: G65016  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-277 <BLAT>  
A:Cross-references: GB:AE000330; GB:U00096; NID:g1788763; PIDN:AAC75477.1; PID:g17887  
A:Experimental source: strain K-12, substrain MG1655  
R:Hryniewicz, M.; Sirko, A.; Paluch, A.; Boeck, A.; Hulanicka, D.  
J. Bacteriol. 172, 3358-3366, 1990  
A:Title: Sulfate and thiosulfate transport in *Escherichia coli* K-12: identification o  
A:Reference number: A35403; MUID:90264335  
A:Accession: B35403  
A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA  
 A:Residues: 1-126,'F',128-133 <HRY>  
 C:Comment: This is one of the membrane-associated components of the binding protein-dep  
 C:Genetics:  
 A:Gene: cysU; cyst  
 A:Map position: 52 min  
 C:Superfamily: maltose transport protein malG  
 C:Keywords: binding protein-dependent transport system; inner membrane; membrane protein

Query Match 65.5%; Score 38; DB 1; Length 277;  
 Best Local Similarity 62.5%; Pred. No. 73;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TWGQYMAV 9  
 : | | | | |  
 Db 43 SMAQYMEV 50

## RESULT 25

B85885  
 hypothetical protein cysU [imported] - Escherichia coli (strain O157:H7)  
 C:Species: Escherichia coli  
 C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
 C:Accession: B85885  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoustis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: B85885  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-277 <STO>  
 A:Cross-references: GB:AE005174; NID:912516799; PIDN:AAG57542.1; GSPDB:GN00145; UMG:Z36  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: cysU  
 C:Superfamily: maltose transport protein malG

Query Match 65.5%; Score 38; DB 2; Length 277;  
 Best Local Similarity 62.5%; Pred. No. 73;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TWGQYMAV 9  
 : | | | | |  
 Db 43 SMAQYMEV 50

## RESULT 26

C86097  
 4-hydroxybenzoate-octaprenyltransferase [imported] - Escherichia coli (strain O157:H7)  
 C:Species: Escherichia coli  
 C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
 C:Accession: C86097  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoustis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: C86097  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-290 <STO>  
 A:Cross-references: GB:AE005174; NID:912518989; PIDN:AAG59239.1; GSPDB:GN00145; UMG:Z56  
 C:Genetics:  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Gene:  
 A:Gene: ubiA  
 C:Superfamily: 4-hydroxybenzoate octaprenyltransferase

Query Match 65.5%; Score 38; DB 2; Length 290;

Best Local Similarity 57.1%; Pred. No. 76;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 OY 3 WGYWWSI 243  
 : | | | | |  
 Db 237 WGYWWSI 243

## RESULT 27

JC2316  
 4-hydroxybenzoate octaprenyltransferase (EC 2.5.1.-) - Escherichia coli  
 C:Species: Escherichia coli  
 C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 20-Apr-2000  
 C:Accession: JC2316; S24361; B42956; S25661; S31432; PC1295; I70801; G65211  
 R:Suzuki, K.; Ueda, M.; Yuasa, M.; Nakagawa, T.; Kawamukai, M.; Matsuda, H.  
 Biosci. Biotechnol. Biochem. 58, 1814-1819, 1994  
 A:Title: Evidence that Escherichia coli ubiA product is a functional homolog of yeast  
 A:Reference number: JC2316; MUID:95072311  
 A:Accession: JC2316  
 A:Molecule type: DNA  
 A:Residues: 1-290 <SUZ>  
 R:Stiebert, M.; Bechtold, A.; Melzer, M.; May, U.; Berger, U.; Schroeder, G.; Schroed  
 FEBS Lett. 307, 347-350, 1992  
 A:Title: Ubiquinone biosynthesis. Cloning of the genes coding for chorismate pyruvate  
 A:Reference number: S24360; MUID:92354744  
 A:Accession: S24361  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-290 <STO>  
 A:Cross-references: EMBL:X66619; NID:943230; PIDN:CAA47182.1; PID:943232  
 R:Nichols, B.P.; Green, J.M.  
 J. Bacteriol. 174, 5309-5316, 1992  
 A:Title: Cloning and sequencing of Escherichia coli ubiC and purification of chorisma  
 A:Reference number: A42956; MUID:9235505  
 A:Accession: B42956  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-290 <NIS>  
 A:Cross-references: GB:M93413; NID:g148106; PIDN:AAA24717.1; PID:g148108  
 R:Nishimura, K.; Nakahigashi, K.; Inokuchi, H.  
 J. Bacteriol. 174, 5762, 1992  
 A:Title: Location of the ubiA gene on the physical map of Escherichia coli.  
 A:Reference number: S25660; MUID:92380960  
 A:Accession: S25661  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-290 <NIS>  
 A:Cross-references: EMBL:X57434; NID:943233; PIDN:CAA40682.1; PID:943235  
 A:Experimental source: strain K12 W3110  
 A>Note: The nucleotide sequence was submitted to the EMBL Data Library, January 1991  
 R:Wolter, F.  
 submitted to the EMBL Data Library, November 1992  
 A:Reference number: S31432  
 A:Accession: S31432  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-290 <WOL>  
 A:Cross-references: EMBL:X69522; NID:941180; PIDN:CAA49270.1; PID:941181  
 R:Lilley, P.E.; Stamford, N.P.D.; Vasudevan, S.G.; Dixon, N.E.  
 Gene 129, 9-16, 1993  
 A:Title: The 92-min region of the Escherichia coli chromosome: location and cloning o  
 A:Reference number: PC1295; MUID:93328130  
 A:Accession: PC1295  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 48-290 <LIL>  
 R:Wu, G.; Williams, H.D.; Gibson, F.; Poole, R.K.  
 J. Gen. Microbiol. 139, 1795-1805, 1993  
 A:Title: Mutants of Escherichia coli affected in respiration: the cloning and nucleot  
 A:Reference number: I55717; MUID:94014977  
 A:Accession: I70801  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-81 <RES>  
 A:Cross-references: GB:M96268; NID:9347886; PIDN:AAA17028.1; PID:9347889  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.F.; Burland, V.; Riley, M.; Co  
 A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of *Escherichia coli* K-12.  
 A:Reference number: A64720; MUID:97426617  
 A:Accession: G65211  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-290 <BLAT>  
 A:Cross-references: GB:AE000477; GB:U00096; NID:92367338; PIDN:AACT7010.1; PID:91790473;  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Genetics:  
 A:Gene: ubiA; cyr  
 A:Map position: 92 min  
 C:Function:  
 A:Description: catalyzes a key step in ubiquinone biosynthesis (condensation of octapren  
 C:Superfamily: 4-hydroxybenzoate octaprenyltransferase  
 C:Keywords: transferase

Query Match 65.5%; Score 38; DB 2; Length 290;  
 Best Local Similarity 57.1%; Pred. No. 76;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 TWGQYMA 9  
 |||||  
 DB 237 WGYWMSI 243

RESULT 28  
 T22928  
 hypothetical protein F58G1.7 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T22928  
 R:Smyle, R.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z19639  
 A:Accession: T22928  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-569 <WIL>  
 A:Cross-references: EMBL:Z81556; PIDN:CA04519.1; GSPDB:GN00020; CESP:F58G1.7  
 A:Experimental source: clone F58G1  
 C:Genetics:  
 A:Gene: CESP:F58G1.7  
 A:Map position: 2  
 A:Introns: 49/3; 126/1; 194/3; 250/3; 401/2; 495/3

Query Match 65.5%; Score 38; DB 2; Length 569;  
 Best Local Similarity 71.4%; Pred. No. 14e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 TWGQYMA 8  
 |||||  
 DB 343 TWEEYMA 349

RESULT 29  
 H64447  
 hypothetical protein MJ1185 - *Methanococcus jannaschii*  
 C:Species: *Methanococcus jannaschii*  
 C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
 C:Accession: H64447  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
 R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
 Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C

A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannasc*  
 A:Reference number: A64300; MUID:96337999  
 A:Accession: H64447  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-622 <BUL>  
 A:Cross-references: GB:U67559; GB:L77117; NID:91591798; PIDN:AAB99186.1; PID:91591812  
 C:Genetics:  
 A:Map position: REV1125034-1123166

Query Match 65.5%; Score 38; DB 2; Length 622;  
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 TWGQYMA 7  
 |||||  
 DB 24 TWGLYMA 29

RESULT 30  
 B64567  
 cytochrome c biogenesis protein - *Helicobacter pylori* (strain 26695)  
 C:Species: *Helicobacter pylori*  
 C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
 C:Accession: B64567  
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R  
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKe  
 son, T.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney,  
 Nature 388, 539-547, 1997  
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,  
 A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
 A:Reference number: A64520; MUID:97394467  
 A:Accession: B64567  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-936 <TOM>  
 A:Cross-references: GB:AE000554; GB:AE000511; NID:92313475; PIDN:AAD07446.1; PID:9231

Query Match 65.5%; Score 38; DB 2; Length 936;  
 Best Local Similarity 57.1%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMA 7  
 ::|||  
 DB 830 ESWGRYMA 836

Search completed: July 3, 2001, 10:23:12  
 Job time: 1337 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: July 3, 2001, 10:23:15 ; Search time 14.6 Seconds  
(without alignments)  
21.116 Million cell updates/sec

Title: US-09-214-836-1  
Perfect score: 58  
Sequence: 1 KTMGQYMAV 9

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	53	91.4	PM17_HUMAN	P40967 homo sapien
2	49	84.5	PM17_MOUSE	O60696 mus musculu
3	44	75.9	P115_CHICK	O98917 gallus galli
4	42	72.4	SLEB_BACSU	P50739 bacillus su
5	42	72.4	CCSA_CHLVV	P56315 chlorella v
6	41	70.7	CCSA_ARATH	P56770 arabidopsis
7	40	69.0	YFAQ_ECOLI	P76463 escherichia
8	40	69.0	YCD5_ECOLI	P75907 escherichia
9	39	67.2	Y090_MYCTU	P25044 saccharomyc
10	39	67.2	Y090_MYCTU	P25044 saccharomyc
11	39	67.2	Y090_MYCTU	P25044 saccharomyc
12	39	67.2	Y090_MYCTU	P25044 saccharomyc
13	38	65.5	PT2B_ARATH	P46032 arabidopsis
14	38	65.5	PT2B_ARATH	P46032 arabidopsis
15	37	63.8	UBIA_ECOLI	P16701 escherichia
16	37	63.8	XYNC_STRLI	P26601 escherichia
17	37	63.8	CCSA_CYACA	P26220 streptomyce
18	37	63.8	CCSA_CYACA	P26220 streptomyce
19	37	63.8	CCSA_CYACA	P26220 streptomyce
20	37	63.8	CCSA_CYACA	P26220 streptomyce
21	37	63.8	CCSA_CYACA	P26220 streptomyce
22	37	63.8	CCSA_CYACA	P26220 streptomyce
23	37	63.8	CCSA_CYACA	P26220 streptomyce
24	37	63.8	CCSA_CYACA	P26220 streptomyce
25	37	63.8	CCSA_CYACA	P26220 streptomyce
26	37	63.8	CCSA_CYACA	P26220 streptomyce
27	37	63.8	CCSA_CYACA	P26220 streptomyce
28	37	63.8	CCSA_CYACA	P26220 streptomyce
29	37	63.8	CCSA_CYACA	P26220 streptomyce
30	37	63.8	CCSA_CYACA	P26220 streptomyce
31	37	63.8	CCSA_CYACA	P26220 streptomyce
32	37	63.8	CCSA_CYACA	P26220 streptomyce
33	37	63.8	CCSA_CYACA	P26220 streptomyce

34	36	62.1	233	1	XYN2_MAGGR	P55335 magnaporthe
35	36	62.1	273	1	OPSR_CANFA	O18914 canis famli
36	36	62.1	301	1	CCSA_GUTH	P22534 guillardi
37	36	62.1	313	1	CCSA_TOBAC	P12216 nicotiana t
38	36	62.1	320	1	CCSA_MARPO	P12214 marichanta
39	36	62.1	321	1	CCSA_MARZE	P46659 zea mays (m
40	36	62.1	321	1	CCSA_ORYSA	P12215 oryza sativ
41	36	62.1	463	1	SYG_MYCTU	O65932 mycobacteri
42	36	62.1	503	1	ALG2_YEAST	P43636 saccharomyc
43	36	62.1	602	1	GAP1_YEAST	P19145 saccharomyc
44	36	62.1	736	1	VM1_REOVL	P12418 reovirus (t
45	36	62.1	736	1	VM1_REOVL	O00335 reovirus (t

## ALIGNMENTS

RESULT	ID	PM17_HUMAN	STANDARD	PRT	661 AA.
AC	P40967	Q16565; Q14817; Q12763; Q14448;			
DT	01-FEB-1995	(Rel. 31, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	MELANOCYTE PROTEIN PMEL 17 PRECURSOR (MELANOCYTE LINEAGE-SPECIFIC				
DE	ANTIGEN GP100) (MELANOMA-ASSOCIATED ME20 ANTIGEN) (ME20/M/ME20S)				
DE	(ME20-M/ME20-S) (95 KDA MELANOCYTE-SPECIFIC SECRETED GLYCOPROTEIN).				
GN	SILV OR PMEL17 OR DI2553E.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID:9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92021023; PubMed=1924386;				
RA	Kwon B.S., Chintamani C., Kozak C.A., Copeland N.G.,				
RA	Gilbert D.J., Jenkins N.A., Barton D., Francke U., Kobayashi Y.,				
RA	Kim K.-K.;				
RT	"A melanocyte-specific gene, Pmel 17, maps near the silver coat color				
RT	locus on mouse chromosome 10 and is in a syntenic region on human				
RT	chromosome 12.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 88:9228-9232(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=94327568; PubMed=7519602;				
RA	Adema G.J., de Boer A.J., Vogel A.M., Loenen W.A., Fijdor C.G.;				
RA	"Molecular characterization of the melanocyte lineage-specific				
RT	antigen gp100.";				
RL	J. Biol. Chem. 269:20126-20133(1994).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96154052; PubMed=8592076;				
RA	Baillin T., Lee S.T., Spritz R.A.;				
RA	"Genomic organization and sequence of DI2553E (Pmel 17), the human				
RT	homologue of the mouse silver (sl) locus.";				
RL	J. Invest. Dermatol. 106:24-27(1996).				
RN	[4]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 25-53.				
RX	MEDLINE=94235165; PubMed=8179825;				
RA	Marresh G.A., Marken J.S., Neubauer M., Aruffo A., Hellstrom I.,				
RA	Hellstrom K.E., Marguardt H.;				
RT	"Cloning and expression of the gene for the melanoma-associated ME20				
RT	antigen.";				
RL	DNA Cell Biol. 13:87-95(1994).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RA	Kwon B.S., Kim K., Heng H.H., Shi X.M., Tsui L., Lee Z.H.,				
RA	Xoun B., Pickard R.T.;				
RP	Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.				
RN	[6]				
RP	SEQUENCE FROM N.A.				
RA	Vogel A.;				
RL	Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.				

	-	FUNCTION:	COULD BE A MELANOGENIC ENZYME. COULD REPRESENT AN ONCOFETAL SELF-ANTIGEN THAT IS NORMALLY EXPRESSED AT LOW LEVELS IN QUIESCENT ADULT MELANOCYTES BUT OVEREXPRESSED BY PROLIFERATING NEONATAL MELANOCYTES AND DURING TUMOR GROWTH. RELEASE OF THE SOLUBLE FORM, ME20-S, COULD PROTECT TUMOR CELLS FROM ANTIBODY MEDIATED IMMUNITY.
CC	-	SUBCELLULAR LOCATION:	TYPE I MEMBRANE PROTEIN (POTENTIAL). THERE IS ALSO A SECRETED SOLUBLE FORM, ME20-S, PROBABLY PRODUCT OF PROTEOLYTIC CLEAVAGE.
CC	-	TISSUE SPECIFICITY:	PREFERENTIALLY EXPRESSED IN MELANOMAS. SOME EXPRESSION WAS FOUND IN DYSPLASTIC NEVI. NOT FOUND IN NORMAL TISSUES NOR IN CARCINOMAS.
CC	-	SIMILARITY:	BELONGS TO THE PMEL-17/MMB FAMILY.
CC	-	SIMILARITY:	CONTAINS 1 PKD DOMAIN.
CC	-	This SWISS-prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to licenses@isb-sib.ch).	
CC		EMBL; M7348;	AAA60121.1; -
DR	EMBL;	S73003;	AAC60634.1; -
DR	EMBL;	U31799;	AAB00386.1; -
DR	EMBL;	U31808;	AAB00386.1; JOINED.
DR	EMBL;	U31807;	AAB00386.1; JOINED.
DR	EMBL;	U31797;	AAB00386.1; JOINED.
DR	EMBL;	U31798;	AAB00386.1; JOINED.
DR	EMBL;	U01874;	AAAI8479.1; -
DR	EMBL;	U20093;	AAI9181.1; -
DR	EMBL;	U19491;	AAI9181.1; JOINED.
DR	EMBL;	M32295;	AAA35930.1; ALT_INIT.
DR	MIM;	155550;	-
DR	InterPro;	IPIR000601;	-
DR	PFam;	PF00801;	PKD. 1.
DR	Prosite;	PS50093;	PKD. 1.
KW	Transmembrane; Glycoprotein;	Signal; Melanin biosynthesis; Repeat; Antigen.	
FT	CHAIN	1	24
FT	DNAIN	25	661
FT	TRANSMEM	25	595
FT	DNAIN	596	616
FT	DNAIN	617	661
FT	DNAIN	255	292
FT	DNAIN	315	444
FT	REPEAT	315	327
FT	REPEAT	328	340
FT	REPEAT	341	353
FT	REPEAT	354	366
FT	REPEAT	367	379
FT	REPEAT	380	392
FT	REPEAT	393	405
FT	REPEAT	406	418
FT	REPEAT	419	431
FT	REPEAT	432	444
FT	CARBOHYD	81	81
FT	CARBOHYD	106	106
FT	CARBOHYD	111	111
FT	CARBOHYD	321	321
FT	CARBOHYD	568	568
FT	CONFLICT	274	274
FT	CONFLICT	587	587
FT	CONFLICT	592	592
FT	CONFLICT	597	597
FT	CONFLICT	642	661
SEQ	SEQUENCE	661 AA;	70255 MM;
			8A904FAFI6715653 CRC64;
	Query Match	91.4%;	Score 53; DB 1; Length 661;
	Best Local Similarity	88.9%;	Pred. No. 0.34;

Matches	8:	Conservative	0:	Mismatches	1:	Indels	0:	Gaps	0:
OY	1	KTWGQYWAY	9		1				
DB	154	KTWGQYWOY	162						
RESULT	2								
PM17_MOUSE									
ID	PM17_MOUSE	STANDARD:	PRT:	626 AA.					
AC	Q60696:								
DT	01-NOV-1997 (Rel. 35, Created)								
DT	01-NOV-1997 (Rel. 35, Last sequence update)								
DT	01-OCT-2000 (Rel. 40, Last annotation update)								
DE	MELANOCTE PROTEIN PMEL 17 PRECURSOR (SILVER LOCUS PROTEIN).								
GN	SILV OR PMEL17 OR D10H12553E OR SI.								
OS	Mus musculus (Mouse).								
OC	Eukaryota; Metazoa; Chordata; Vertebrate; Euteleostomi;								
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
OX	NCBI_Taxid=10090;								
RA	SEQUENCE FROM N.A.								
RC	STRAIN=C57BL/6; TISSUE=Skin;								
RX	MEDLINE=95175358; Pubmed=7870580;								
RA	Kwon B.S., Halabaz R., Ponnazhagan S., Kim K., Chintamani C.,								
RT	Bennett D., Pickard R.T.;								
RT	"Mouse silver mutation is caused by a single base insertion in the								
RL	putative cytoplasmic domain of Pmel 17.";								
CC	Nucleic Acids Res. 23:154-158(1995).								
CC	-1- FUNCTION: COULD BE A MELANOGENIC ENZYME.								
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).								
CC	-1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCTES.								
CC	-1- DISEASE: DEFECTS IN PMEL17 ARE THE CAUSE OF THE SILVER COAT COLOR								
CC	WHICH SEEMS TO BE DUE TO PREMATURE DEATH OF PIGMENT CELLS DURING								
CC	THE HAIR CYCLE.								
CC	-1- SIMILARITY: BELONGS TO THE PMEL-17/NUB FAMILY.								
CC	-1- SIMILARITY: CONTAINS 1 PKD DOMAIN.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>								
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).								
CC	-----								
DR	EMBL, U14133; AAA69538.1; .								
DR	MCD; MGI:98301; sl.								
DR	PROSITE, PS50093; PKD; 1.								
KW	Transmembrane; Glycoprotein; Signal; Melanin biosynthesis; Repeat;								
KW	disease mutation.								
FT	SIGNAL	1	24						
FT	CHAIN	25	626						
FT	DOMAIN	25	562						
FT	TRANSMEM	563	583						
FT	DOMAIN	584	626						
FT	DOMAIN	255	292						
FT	DOMAIN	315	411						
FT	REPEAT	315	327						
FT	REPEAT	328	340						
FT	REPEAT	341	353						
FT	REPEAT	354	366						
FT	REPEAT	367	379						
FT	REPEAT	380	392						
FT	REPEAT	393	411						
FT	CARBOHYD	81							
FT	CARBOHYD	106	106						
FT	CARBOHYD	111	111						
FT	CARBOHYD	535	535						
FT	CARBOHYD	170	170						
FT	VARIANT	1							

FT VARIANT 603 626 AAPASGLRANGIENSPLSGOQV -> SSASLRSSRPV  
 FT SEQUENCE 626 AA; 65980 MM; 7AB941D2E3FB1044 CRC64;  
 SQ

Query Match 84.5%; Score 49; DB 1; Length 626;  
 Best Local Similarity 77.8%; Pred. No. 1.3;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTMGOYNAV 9  
 DB 154 KTMGWYQV 162

RESULT 3  
 ID P115\_CHICK STANDARD; PRT: 762 AA.  
 AC Q98917;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE MELANOSOMAL MATRIX PROTEIN 115 KDA PROTEIN PRECURSOR.  
 GN MFPI13.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WHITE LECHORN; TISSUE-Retinal pigment epithelium;  
 RX MEDLINE=92020667; PubMed=1924173;  
 RA Mochi M., Agata K., Eguchi G.;  
 RT "Complete sequence and expression of a cDNA encoding a chicken  
 RT 115-kDa melanosomal matrix protein.";  
 RL Pigment Cell Res. 4:41-47(1991).  
 RN [2]  
 RP CHARACTERIZATION.  
 RC STRAIN-WHITE LECHORN; TISSUE-Retinal pigment epithelium;  
 RX MEDLINE=88311098; PubMed=3409326;  
 RA Mochi M., Agata K., Kobayashi H., Yamamoto T.S., Eguchi G.;  
 RT "Expression of gene coding for a melanosomal matrix protein  
 RT transcriptionally regulated in the transdifferentiation of chick  
 RT embryo pigmented epithelial cells.";  
 RL Cell Differ. 24:67-74(1988).

CC -1- FUNCTION: MIGHT BE REQUIRED FOR POLYMERIZATION OF MELANIN ONTO THE  
 CC CORE STRUCTURE OF MELANOSOMES WITH ENZYMIC FUNCTION OF TYROSINASE.  
 CC -1- SUBCELLULAR LOCATION: ON THE FIBROUS MATRIX STRUCTURE OF THE  
 CC PREMELANOSOME.  
 CC -1- TISSUE SPECIFICITY: SPECIFIC TO PIGMENTED EPITHELIAL CELLS AND  
 CC MELANOCYTES. NOT EXPRESSED IN LENS, NEURAL RETINA, BRAIN, HEART,  
 CC GIZZARD OR LIVER.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING THE REDIFFERENTIATION OF  
 CC PIGMENTED EPITHELIAL CELLS (PEC).  
 CC -1- PTM: GLYCOSYLATED.  
 CC -1- SIMILARITY: BELONGS TO THE PMEL-17/NMB FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 PKD DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; D88348; BAAL3589.1; -;  
 DR InterPro; IPR000601; -;  
 DR Pfam; PF00801; PKD. 1.  
 DR PROSITE; PSS0093; PKD. 1.  
 KM Signal; Glycoprotein; Repeat.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 762 MELANOSOMAL MATRIX PROTEIN 115 KDA

FT FT  
 FT DOMAIN 223 323 PROTEIN.  
 FT PKD.  
 FT 4 X 20-24 AA APPROXIMATE TANDEM REPEATS.  
 FT REPEAT 441 532  
 FT REPEAT 441 464  
 FT REPEAT 465 488  
 FT REPEAT 489 508  
 FT REPEAT 509 532  
 FT CARBOHYD 111 111  
 FT CARBOHYD 115 115  
 FT CARBOHYD 115 115  
 FT CARBOHYD 346 346  
 FT CARBOHYD 651 651  
 FT CARBOHYD 659 659  
 FT CARBOHYD 659 659  
 SQ SEQUENCE 762 AA; 77356 MM; 172C8DB4FDC766 CRC64;

Query Match 75.9%; Score 44; DB 1; Length 762;  
 Best Local Similarity 75.0%; Pred. No. 8.9;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 OY 2 TWGOYNAV 9  
 DB 161 TWGRTWQV 168

RESULT 4  
 ID SLEB\_BACSU STANDARD; PRT: 305 AA.  
 AC P50739;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE SPORE-CORTEX-LYTIC ENZYME PRECURSOR.  
 GN SLEB.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=96427343; PubMed=8830707;  
 RA Moriyama R., Hattori A., Miyata S., Makino S.;  
 RT "A gene (sleB) encoding a spore cortex-lytic enzyme from Bacillus  
 RT subtilis and response of the enzyme to L-alanine-mediated  
 RT germination.";  
 RL J. Bacteriol. 178:6059-6063(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / MARBURG;  
 RX MEDLINE=96349105; PubMed=8760912;  
 RA Sorokin A.V., Azevedo V., Zumbstein E., Galleron N., Ehrlich S.D.,  
 RA Serror P.;  
 RT "Sequence analysis of the Bacillus subtilis chromosome region between  
 RT the sera and kds loci cloned in a yeast artificial chromosome.";  
 RL Microbiology 142:2005-2016(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; D79978; BAAL1473.1; -;  
 DR EMBL; L47648; AAC83957.1; -;  
 DR EMBL; Z99115; CAB14209.1; -;  
 DR EMBL; Z99116; CAB14225.1; -;  
 DR Subtilist; BG11439; SLEB.  
 DR InterPro; IPR002477; -;  
 DR Pfam; PF01471; PG\_Binding\_1; 1.  
 KM Signal.  
 FT SIGNAL 1 29 POTENTIAL.

FT CHAIN 30 305 SPORE-CORTEX-LYTIC ENZYME.  
SQ SEQUENCE 305 AA; 34001 MW; 9DF1305975F5BE16 CRC64;

Query Match 72.4%; Score 42; DB 1; Length 305;  
Best Local Similarity 85.7%; Pred. No. 7.8;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 MCGYNAV 9  
11111  
DB 67 MCGYNAV 73

RESULT 5  
ID CCSA\_CHLVU STANDARD; PRT; 315 AA.  
AC P56315;

DT 15-JUL-1998 (rel. 36, Created)  
DT 15-JUL-1998 (rel. 36, Last sequence update)  
DT 15-JUL-1998 (rel. 36, Last annotation update)  
DE CYTOCHROME C BIOGENESIS PROTEIN CCSA.

OS Chloroplast.

OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;  
OC Chlorellaceae; Chlorella.

OX NCBI\_TaxID=3077;

[1]

SEQUENCE FROM N.A.

RC STRAIN-IAM C-27 / TAMIVA;

RA MEDLINE-97303241; PubMed-9159184;

RA Waksugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,  
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,  
RA Inamura A., Yoshinaga K., Sugita M.;

\*Complete nucleotide sequence of the chloroplast genome from the

RT green alga *Chlorella vulgaris*: the existence of genes possibly  
involved in chloroplast division\*;

RT Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).

-1- FUNCTION: REQUIRED DURING CYTOCHROME BIOGENESIS AT THE STEP OF

HEME ATTACHMENT (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO THE CCMF/CYCK/CCL1/NRFE/CCSA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AB001684; BAA57962.1; -.

DR InterPro; IPR002541; -.

DR Pfam; PF01578; CYTC\_asm; 1.

KW Cytochrome c-type biogenesis; Chloroplast.

SQ SEQUENCE 315 AA; 35471 MW; 5020388E94FEAF10 CRC64;

Query Match 72.4%; Score 42; DB 1; Length 315;  
Best Local Similarity 62.5%; Pred. No. 8;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYWA 8  
11111  
DB 243 ETWGWYWS 250

RESULT 6  
ID CCSA\_ARATH STANDARD; PRT; 328 AA.  
AC P56770;

DT 30-MAY-2000 (rel. 39, Created)

DT 30-MAY-2000 (rel. 39, Last sequence update)  
DT 01-OCT-2000 (rel. 40, Last annotation update)

DE CYTOCHROME C BIOGENESIS PROTEIN CCSA.  
GN CCSA.

OS Arabidopsis thaliana (Mouse-ear cress).

OG Chloroplast.

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

OC Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;

[1]

SEQUENCE FROM N.A.

RC STRAIN-CV. COLUMBIA;

RA MEDLINE-20039611; PubMed-10574454;

RA Sato S., Nakamura Y., Kaneko T., Asamizu E., Tabata S.;

RT \*Complete structure of the chloroplast genome of *Arabidopsis*  
thaliana\*.

RT DNA Res. 6:283-290(1999).

-1- FUNCTION: REQUIRED DURING CYTOCHROME BIOGENESIS AT THE STEP OF

HEME ATTACHMENT (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO THE CCMF/CYCK/CCL1/NRFE/CCSA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AP000423; BAA84436.1; -.

DR InterPro; IPR002541; -.

DR Pfam; PF01578; CYTC\_asm; 1.

KW Cytochrome c-type biogenesis; Chloroplast.

SQ SEQUENCE 328 AA; 37732 MW; C8BD1508E2924D6F CRC64;

Query Match 70.7%; Score 41; DB 1; Length 328;  
Best Local Similarity 71.4%; Pred. No. 12;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYWA 7  
11111  
DB 255 ETWGSYWM 261

RESULT 7  
ID YFAO\_ECOLI STANDARD; PRT; 549 AA.  
AC P76463;

DT 01-OCT-2000 (rel. 40, Created)

DT 01-OCT-2000 (rel. 40, Last sequence update)

DT 01-OCT-2000 (rel. 40, Last annotation update)

DE HYPOTHETICAL 61.5 KDA PROTEIN IN ATOB-GYRA INTERGENIC REGION

DE PRECURSOR.

GN YFAO.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI\_TaxID=562;

[1]

SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;

RA MEDLINE-97426617; PubMed-9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;

\*The complete genome sequence of *Escherichia coli* K-12\*.

Science 277:1453-1474(1997).

RT

RL

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way



modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).)  
 CC -----  
 DR EMBL; AE000312; AAC75286.1; -  
 DR EcGene; EG14079; yfao.  
 KW Hypothetical protein; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 549  
 FT SEQUENCE 549 AA; 61475 MW; 72C26716D953C9D1 CRC64;

Query Match 69.0%; Score 40; DB 1; Length 549;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGYW 7  
 |||||  
 Db 125 WGYW 129

RESULT 8  
 ID YCDS\_ECOLI STANDARD; PRT; 807 AA.  
 AC P75907;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE HYPOTHETICAL 92.2 KDA PROTEIN IN PHOH-CSSG INTERGENIC REGION  
 DE PRECURSOR.  
 GN YCDS.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subphylum; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE=97436617; PubMed=9276503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=97061202; PubMed=8905232;  
 RA Oshima T., Alha H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizoduchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horiuchi T.;  
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 12.7-28.0 min region on the linkage map.";  
 RL DNA Res. 3:137-155(1996).  
 CC -1 SUBCELLULAR LOCATION: OUTER MEMBRANE (POTENTIAL).  
 CC -1 SIMILARITY: STRONG. TO Y. PESTIS HEMIN-BINDING PROTEIN HNSH.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).)  
 CC -----  
 DR EMBL; AE000204; AAC74109.1; -  
 DR EMBL; D90739; BAA35806.1; -  
 DR EMBL; D90740; BAA35809.1; -  
 DR EcGene; EG13865; ycds.

KW Hypothetical protein; Outer membrane; Signal.  
 FT SIGNAL 1 26  
 FT CHAIN 27 807  
 FT SEQUENCE 807 AA; 92207 MW; B20067C3D41723FD CRC64;

Query Match 69.0%; Score 40; DB 1; Length 807;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGYW 7  
 |||||  
 Db 314 WGYW 318

RESULT 9  
 ID FCP\_ISOGA STANDARD; PRT; 208 AA.  
 AC Q39709;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE FUCOXANTHIN-CHLOROPHYLL A-C BINDING PROTEIN, CHLOROPLAST PRECURSOR  
 DE (FCP).  
 GN FCP.  
 OS Isochrysis galbana.  
 OC Eukaryota; Haptophyceae; Isochrysidales; Isochrysis.  
 OX NCBI\_TaxID=37099;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN-DUN;  
 RX MEDLINE=94325461; PubMed=8049362;  
 RA Laroche J., Henry D., Wyman K., Suenik A., Falkowski P.;  
 RT "Cloning and nucleotide sequence of a cDNA encoding a major  
 RT fucoxanthin-chlorophyll a/c-containing protein from the chrysophyte  
 RT Isochrysis galbana: implications for evolution of the cab gene  
 RT family.";  
 RL Plant Mol. Biol. 25:355-368(1994).  
 CC -1 FUNCTION: THE LIGHT-HARVESTING COMPLEX (LHC) FUNCTIONS AS A LIGHT  
 CC RECEPTOR. IT CAPTURES & DELIVERS EXCITATION ENERGY TO PHOTOSYSTEMS  
 CC WITH WHICH IT IS CLOSELY ASSOCIATED. ENERGY IS TRANSFERRED FROM  
 CC THE CAROTENOID AND CHL C (OR B) TO CHL A AND THE PHOTOSYNTHETIC  
 CC REACTION CENTERS WHERE IT IS USED TO SYNTHESIZE ATP AND REDUCING  
 CC POWER.  
 CC -1 SUBUNIT: THE LHC COMPLEX OF CHROMOPHYTIC ALGAE IS COMPOSED OF  
 CC FUCOXANTHIN, CHLOROPHYLL A AND C BOUND NON-COVALENTLY BY  
 CC PIGMENTS IN LHC; FUCOXANTHIN: CHLOROPHYLL C: CHLOROPHYLL A IS  
 CC (0.6-1): (0.1-0.3): (1).  
 CC -1 SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE. FCPs ARE  
 CC PROBABLY TRANSPORTED ACROSS THE ENDOPLASMIC RETICULUM MEMBRANE  
 CC THAT SURROUND THE PLASTID VIA A SIGNAL PEPTIDE, FOLLOWED BY  
 CC TRANSLOCATION ACROSS THE THYLAKOID MEMBRANE VIA A TRANSIT PEPTIDE.  
 CC -1 INDUCTION: EXPRESSION IS INCREASED 5-FOLD UNDER CONDITIONS OF  
 CC LOW LIGHT.  
 CC -1 SIMILARITY: BELONGS TO THE FCP FAMILY OF LIGHT-HARVESTING  
 CC PROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).)  
 CC -----  
 DR EMBL; X77333; CAA54547.1; -  
 DR InterPro; IPR001344; -  
 DR Pfam; PF00504; chloroa\_b-bind; 2  
 DR Light-harvesting polypeptide; Chloroplast; Photosynthesis;  
 KW Photosystem II; Multigene family; Chlorophyll; Transmembrane;  
 KW Thylakoid membrane; Transit peptide.  
 FT TRANSIT 1 31  
 FT CHLOROPLAST (PROBABLE).

FT CHAIN 32 208 FUCOXANTHIN-CHLOROPHYLL A-C BINDING  
 FT TRANSMEM 102 118 POTENTIAL.  
 SO SEQUENCE 208 AA; 22471 MW; 21A36700137A0F1B CRC64;

Query Match 67.2%; Score 39; DB 1; Length 208;  
 Best Local Similarity 83.3%; Pred. No. 16;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 TMGOY 7  
 DB 198 TMGOY 203

RESULT 10  
 Y090 MYCTU  
 ID Y090 MYCTU STANDARD; PRT; 256 AA.

AC 010887;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE HYPOTHEICAL 27.8 KDA PROTEIN RV0090.  
 GN RV0090 OR MTCY251.08.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

NCBI\_TaxID=1773;  
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 Gordon S.V., Eigler R., Gars S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osbourne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,  
 RA Taylor K., Whitehead S., Barrell B.G.;  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: Z74410; CAA8926.1; -

DR Tuberculin: RV0090; -

KW Hypothetical protein; Transmembrane.

FT TRANSMEM 155 175 POTENTIAL.

FT TRANSMEM 203 223 POTENTIAL.

SO SEQUENCE 256 AA; 27837 MW; 01033C21199DEC51 CRC64;

Query Match 67.2%; Score 39; DB 1; Length 256;  
 Best Local Similarity 66.7%; Pred. No. 19;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGOY 9  
 DB 177 KTWGOY 185

RESULT 11  
 PTP1 YEAST  
 ID PTP1 YEAST STANDARD; PRT; 335 AA.  
 AC P25044;

DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE PROTEIN-TYROSINE PHOSPHATASE 1 (EC 3.1.3.48) (PTPase 1).  
 GN PTP1 OR YDL230W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

NCBI\_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=91302312; PubMed=1649172;

RX Guan K., Deschenes R.J., Qiu H., Dixon J.E.;

RT "Cloning and expression of a yeast protein tyrosine phosphatase.";

RL J. Biol. Chem. 266:12964-12970(1991).

RN [2]

RP SEQUENCE FROM N.A.

RA Rasmussen S.W.;

RL Submitted (Jul-1996) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: IS NOT REQUIRED FOR VEGETATIVE GROWTH.

CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =

CC PROTEIN TYROSINE + ORTHOPHOSPHATE.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-

CC TYROSINE PHOSPHATASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M64062; AAA34923.1; -

DR EMBL: Z74278; CAA8809.1; -

DR PIR: A39862; A39862.

DR HSP: P18052; 1YFO.

DR SGP: S0002389; PTP1.

DR InterPro: IPR000242; -

DR InterPro: IPR000387; -

DR Pfam: PF00102; Y-phosphatase; 1.

DR PRINTS: PR00700; PTPPHPTASE.

DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; 1.

DR PROSITE: PS50056; TYR\_PHOSPHATASE\_2; 1.

DR PROSITE: PS50055; TYR\_PHOSPHATASE\_PTP; 1.

KW Hydrolase.

FT ACT\_SITE 252 252 BY SIMILARITY.

SO SEQUENCE 335 AA; 38868 MW; 15F71E50694BE562 CRC64;

Query Match 67.2%; Score 39; DB 1; Length 335;  
 Best Local Similarity 71.4%; Pred. No. 24;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGOY 7  
 DB 108 KTWGOY 114

RESULT 12  
 PTP2B ARATH  
 ID PTP2B ARATH STANDARD; PRT; 585 AA.  
 AC P46032;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE PEPTIDE TRANSPORTER PTP2B (HISTIDINE TRANSPORTING PROTEIN).

GN PTP2B OR NTR1 OR AT2G02040 OR F14H20.11.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

OC Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, LANDSBERG ERECTA;  
 RA Song W., Steiner H.-Y., Zhang L., Naider F., Stacey G.,  
 RL Becker J.M.;  
 RN Submitted (xxx-1995) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, C24;  
 RA MEDLINE=94307379; PubMed=8033999;  
 RA Frommer W.B., Hummel S., Rentsch D.;  
 RT "Cloning of an Arabidopsis histidine transporting protein related to  
 RT nitrate and peptide transporters.";  
 RL FEMS Lett. 347:185-189(1994).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=CV, COLUMBIA;  
 RA MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
 RA Cronin L.A., Shen M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E.,  
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,  
 RA Salzberg S.L., Fraser C.M., Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 402:761-768(1999).  
 CC -1- FUNCTION: PEPTIDE TRANSPORT. HIGH AFFINITY, LOW CAPACITY  
 CC TRANSPORTER. CAN ALSO TRANSPORT HISTIDINE.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE PPR2 FAMILY OF TRANSPORTERS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L39082; AAB00858.1; -;  
 DR EMBL: X77503; CA54634.1; -;  
 DR EMBL: AC006532; AAD20096.1; -;  
 DR InterPro: IPR00109; -;  
 DR Pfam: PF00854; PTR2.1;  
 DR PROSITE: PS01022; PTR2.1;  
 DR PROSITE: PS01023; PTR2.2; 1;  
 KW Peptide transport; Transport; Transmembrane.  
 FT TRANSMEM 91 111 POTENTIAL.  
 FT TRANSMEM 116 136 POTENTIAL.  
 FT TRANSMEM 154 174 POTENTIAL.  
 FT TRANSMEM 200 220 POTENTIAL.  
 FT TRANSMEM 228 248 POTENTIAL.  
 FT TRANSMEM 351 371 POTENTIAL.  
 FT TRANSMEM 387 407 POTENTIAL.  
 FT TRANSMEM 431 451 POTENTIAL.  
 FT TRANSMEM 472 492 POTENTIAL.  
 FT TRANSMEM 511 531 POTENTIAL.  
 FT TRANSMEM 556 576 POTENTIAL.  
 FT CONFLICT 334 334 R -> ED (IN REF. 2).  
 FT SEQUENCE 585 AA; 64421 MM; C58F81947762D97 CMC64;

Query Match 67.2%; Score 39; DB 1; Length 585;  
 Best Local Similarity 57.1%; Pred. No. 40;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Oy 3 WGVYNAV 9  
 Db 110 WGRVWTI 116

RESULT 13  
 ID CYST\_ECOLI STANDARD; PRT; 277 AA.  
 AC P16701;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE SULFATE TRANSPORT SYSTEM PERMEASE PROTEIN CYST.  
 GN CYSU OR CYST.  
 OS Escherichia coli.  
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Escherichia.  
 OX NCBI\_TaxID=562;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=K12;  
 RC MEDLINE=90264334; PubMed=2188958;  
 RA Sirko A., Hryniewicz M.M., Hulanicka D.M., Boeck A.;  
 RT "Sulfate and thiosulfate transport in Escherichia coli K-12:  
 RT nucleotide sequence and expression of the cystWAM gene cluster.";  
 RL J. Bacteriol. 172:3351-3357(1990).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=K12 / MG1655;  
 RC MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=K12;  
 RC MEDLINE=97349980; PubMed=9205837;  
 RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,  
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,  
 RA Mizobuchi K., Mori H., Nakada S., Nakamura Y., Nishimoto H.,  
 RA Oshima T., Oyama S., Saito N., Sempel G., Satoh Y., Sivasubram S.,  
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,  
 RA Yamagata S., Horiuchi T.;  
 RT "Construction of a contiguous 874-kb sequence of the Escherichia coli  
 RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and  
 RT analysis of its sequence features.";  
 RL DNA Res. 4:91-113(1997).  
 CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
 CC FOR SULFATE AND THIOSULFATE. PROBABLY RESPONSIBLE FOR THE  
 CC TRANSLLOCATION OF THE SUBSTRATE ACROSS THE MEMBRANE.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE  
 CC (POTENTIAL).  
 CC -1- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-  
 CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE CYSTW  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M32101; AAA23637.1; -;  
 DR EMBL: AE000330; AAC75477.1; -;  
 DR EMBL: D90871; BAA16298.1; -;  
 DR EMBL: D90872; BAA16307.1; -;  
 DR PIR: A35402; QRECSF.  
 DR PIR: B35403; B35403.  
 DR EcGene: BG10197; cySU.  
 DR InterPro: IPR000515; -;  
 DR Pfam: PF00528; BPD\_transp.1.  
 DR PROSITE: PS00402; BPD\_TRANSP\_INN\_MEMBR.1.

KW Inner membrane; Transmembrane; Sulfate transport; Transport.  
FT TRANSMEM 21 39 POTENTIAL.  
FT TRANSMEM 64 80 POTENTIAL.  
FT TRANSMEM 105 121 POTENTIAL.  
FT TRANSMEM 140 156 POTENTIAL.  
FT TRANSMEM 188 204 POTENTIAL.  
FT TRANSMEM 247 259 POTENTIAL.  
SQ SEQUENCE 277 AA; 30291 MW; 1342821B0DE24459 CRC64.  
[5]  
Query Match 65.5%; Score 38; DB 1; Length 277;  
Best Local Similarity 62.5%; Pred. No. 29;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
OY 2 TWGQYMAV 9  
DB 43 SNAQYMEV 50  
[6]  
RESULT 14  
UBIA.ECOLI STANDARD; PRT; 290 AA.  
AC P26601;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE 4-HYDROXYBENZATE OCTAPRENYLTRANSFERASE (EC 2.5.1.-) (4-HB  
DE POLYPRENYLTRANSFERASE).  
GN UBIA OR CYR.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RX MEDLINE-92355505; PubMed1644758;  
RA Nichols B.P., Green J.M.;  
RT "Cloning and sequencing of Escherichia coli ublc and purification of  
RT chorismate lyase.";  
RL J. Bacteriol. 174:5309-5316(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / MC4100;  
RX MEDLINE-92354744; PubMed1644192;  
RA Siebert M., Bechtold A., Melzer M., May U., Berger U., Schroeder G.,  
RT Schroeder J., Severin K., Heide L.;  
RT "Ubiquinone biosynthesis. Cloning of the genes coding for chorismate  
RT pyruvate-lyase and 4-hydroxybenzoate octaprenyl transferase from  
RT Escherichia coli.";  
RL FEBS Lett. 307:347-350(1992).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / W3110;  
RX MEDLINE-92380960; PubMed1512213;  
RA Nishimura K., Nakahigashi K., Inokuchi H.;  
RT "Location of the ubla gene on the physical map of Escherichia coli.";  
RL J. Bacteriol. 174:5762-5762(1992).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RA Wolter F.P.;  
RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RX MEDLINE-94089392; PubMed8265357;  
RA Biatner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,  
RT Daniels D.L.;  
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the  
RT region from 89.2 to 92.8 minutes.";  
RL Nucleic Acids Res. 21:5408-5417(1993).  
RN [6]

RP SEQUENCE OF 1-80 FROM N.A.  
RC STRAIN-K12;  
RX MEDLINE-94014977; PubMed8409922;  
RA Wu G., Williams H.D., Gibson F., Poole R.K.;  
RT "Mutants of Escherichia coli affected in respiration: the cloning and  
RT nucleotide sequence of ubla, encoding the membrane-bound p-  
RT hydroxybenzoate:octaprenyltransferase.";  
RL J. Gen. Microbiol. 139:1795-1805(1993).  
RN [7]  
RP CHARACTERIZATION.  
RX MEDLINE-94207029; PubMed8155731;  
RA Welter M., Heide L.;  
RT "Characterization of poly(prenyl)phosphate:4-hydroxybenzoate  
RT poly(prenyl)transferase from Escherichia coli.";  
RL Biochim. Biophys. Acta 1212:93-102(1994).  
RN [8]  
RP CHARACTERIZATION.  
RX MEDLINE-95072311; PubMed7765507;  
RA Suzuki K., Ueda M., Yuasa M., Nakagawa T., Kawamukai M., Matsuda H.;  
RT "Evidence that Escherichia coli ubla product is a functional homolog  
RT of yeast COO2, and the regulation of ubla gene expression.";  
RL Biosci. Biotechnol. Biochem. 58:1814-1819(1994).  
CC -1- FUNCTION: SYNTHESIS OF 3-OCTAPRENYL-4-HYDROXYBENZATE.  
CC -1- CATALYTIC ACTIVITY: 4-HYDROXYBENZATE + FARNESYLARNESYLGERANIOL  
CC -1- COFACTOR: REQUIRES MAGNESIUM.  
CC -1- PATHWAY: SECOND STEP IN UBIQUINONE BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.  
CC -1- SIMILARITY: BELONGS TO THE UBIA PRENYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
CC -----  
DR EMBL; M93136; AAA24712.1; -  
DR EMBL; M93413; AAA24717.1; -  
DR EMBL; X66619; CAA47182.1; -  
DR EMBL; X57434; CAA40682.1; -  
DR EMBL; X69522; CAA49270.1; -  
DR EMBL; U00006; AAC43134.1; -  
DR EMBL; AE000477; AAC77010.1; -  
DR EMBL; M96268; AAA17028.1; -  
DR EMBL; X63407; CAA45003.1; -  
DR PIR; BA2956; B42956.  
DR PIR; S24361; S24361.  
DR PIR; S25661; S25661.  
DR PIR; S31432; S31432.  
DR PIR; JC2316; JC2316.  
DR EcoGene; EG11370; ubla.  
DR InterPro; IPR000537; -  
DR Pfam; PF01040; UBIA; 1.  
DR PROSITE; PS00943; UBIA; 1.  
KW Ubiquinone biosynthesis; Transferase; Transmembrane; Inner membrane;  
KW Magnesium.  
FT TRANSMEM 23 43 POTENTIAL.  
FT TRANSMEM 46 66 POTENTIAL.  
FT TRANSMEM 100 120 POTENTIAL.  
FT TRANSMEM 141 161 POTENTIAL.  
FT TRANSMEM 163 183 POTENTIAL.  
FT TRANSMEM 213 233 POTENTIAL.  
FT TRANSMEM 234 254 POTENTIAL.  
FT TRANSMEM 268 288 POTENTIAL.  
SQ SEQUENCE 290 AA; 32511 MW; F10FED1DA30E115 CRC64.

Query Match 65.5%; Score 38; DB 1; Length 290;  
Best Local Similarity 57.1%; Pred. No. 30;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 WGOYMAV 9  
 11111:  
 Db 237 WGYTMSI 243

RESULT 15  
 XINC STRLI STANDARD; PRT: 240 AA.  
 AC P26220:  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE ENDO-1,4-BETA-XYLANASE C PRECURSOR (EC 3.2.1.8) (XYLANASE C)  
 GN (1,4-BETA-D-XYLAN XYLANOXYLASE C).  
 OS Streptomyces lividans.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1916;  
 RN [1]  
 RC SEQUENCE FROM N.A., AND SEQUENCE OF 50-80.  
 RP STRAIN=66 / 1326;  
 RX MEDLINE=92077439; PubMed=1743521;  
 RA Sharek F., Roy C., Yaguchi M., Morosoli R., Kluepfel D.;  
 RT "Sequences of three genes specifying xylanases in Streptomyces  
 lividans.";  
 RL Gene 107:75-82(1991).  
 CC -1- FUNCTION: CONTRIBUTES TO HYDROLASE HEMICELLULOSE. THE MAJOR  
 CC COMPONENT OF PLANT CELL WALLS.  
 CC -1- CATALYTIC ACTIVITY: ENDOMYDOLYSIS OF 1,4-BETA-D-XYLOSIDIC  
 CC LINKAGES IN XYLANS.  
 CC -1- PATHWAY: XYLAN DEGRADATION.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO CELLULOSE FAMILY G (FAMILY 11 OF GLYCOSYL  
 CC HYDROLASES).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M64553; AAA26836.1; -;  
 DR EMBL: A25307; CAA01768.1; -;  
 DR PIR: JS0591; JS0591.  
 DR HSP: P09850; IBCX.  
 DR InterPro: IPR001137; -;  
 DR Pfam: PF00457; Glyco\_hydro\_11; 1.  
 DR PRINTS: PR00911; GLHYDLASE1.  
 DR PROSITE: PS00776; GLYCOSYL\_HYDROL\_F11\_1; 1.  
 DR PROSITE: PS00777; GLYCOSYL\_HYDROL\_F11\_2; 1.  
 KW xylan degradation; Hydrolase; Glycosidase; Signal.  
 FT SIGNAL 1 49  
 FT CHAIN 50 240 ENDO-1,4-BETA-XYLANASE C.  
 FT ACT\_SITE 134 134 NUCLEOPHILE (BY SIMILARITY).  
 FT ACT\_SITE 226 226 PROTON DONOR (BY SIMILARITY).  
 SQ SEQUENCE 240 AA; 25673 MW; FC65415780142CA CRC64;

Query Match 63.8%; Score 37; DB 1; Length 240;  
 Best Local Similarity 66.7%; Pred. No. 36;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGOYMAV 9  
 11111:  
 Db 178 KTFQOYMSV 186

RESULT 16  
 CCSA\_CYACA STANDARD; PRT: 306 AA.  
 ID CCSA\_CYACA

AC P31564;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE CYTOCHROME C BIOGENESIS PROTEIN CCSA.  
 GN CCSA.  
 OS Cyanidium caldarium (Caldieria sulphuraria).  
 OS Chloroplast.  
 OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;  
 OC Cyanidium.  
 OX NCBI\_TaxID=2771;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=14-1-1 / ISOLATE 107 79/GOETTINGEN;  
 RX MEDLINE=93004479; PubMed=1391770;  
 RA Valentin K., Maid U., Emich A., Zetsche K.;  
 RT "Organization and expression of a phycoobiliprotein gene cluster from  
 the unicellular red alga Cyanidium caldarium.";  
 RL Plant Mol. Biol. 20:267-276(1992).  
 CC -1- FUNCTION: REQUIRED DURING CYTOCHROME BIOGENESIS AT THE STEP OF  
 CC HEME ATTACHMENT (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE CCMF/CYCK/CCL/NRPF/CCSA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X57150; CAA0439.1; -;  
 DR PIR: S14520; S14520.  
 DR PIR: S25309; S25309.  
 DR InterPro: IPR002541; -;  
 DR Pfam: PF01578; CytC\_asm; 1.  
 KW Cytochrome c-type biogenesis; Chloroplast.  
 SQ SEQUENCE 306 AA; 35263 MW; 4C3C9AFB48292F4F CRC64;

Query Match 63.8%; Score 37; DB 1; Length 306;  
 Best Local Similarity 50.0%; Pred. No. 45;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTWGOYMA 8  
 11111:  
 Db 235 EAWGYSWS 242

RESULT 17  
 CCSA\_ODOSI STANDARD; PRT: 312 AA.  
 ID CCSA\_ODOSI  
 AC P49523;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE CYTOCHROME C BIOGENESIS PROTEIN CCSA.  
 GN CCSA.  
 OS Odontella sinensis.  
 OS Chloroplast.  
 OC Eukaryota; Stramenopiles; Bacillariophyta; Coscinodiscophyceae;  
 OC Biddulphiophycidae; Eupodiscales; Eupodisaceae; Odontella.  
 OX NCBI\_TaxID=2839;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;  
 RT "The chloroplast genome of a chlorophyll a+c-containing alga,  
 Odontella sinensis.";  
 RL Plant Mol. Biol. Rep. 13:336-342(1995).  
 CC -1- FUNCTION: REQUIRED DURING CYTOCHROME BIOGENESIS AT THE STEP OF  
 CC HEME ATTACHMENT (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE CCMF/CYCK/CCL/NRPF/CCSA FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z67753; CAA91615.1; -  
 DR InterPro: IPR002541; -  
 DR Pfam: PF01578; CytC\_asm.1.  
 KW Cytochrome c-type biogenesis; Chloroplast.  
 SQ SEQUENCE 312 AA; 35436 MW; 748390BE4BE3C5B9 CRC64;

Query Match 63.8%; Score 37; DB 1; Length 312;  
 Best Local Similarity 50.0%; Pred. No. 46;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTMGOYWA 8  
 : || ||:  
 Db 241 EAMGSYMS 248

RESULT 18  
 MANA\_STRMU STANDARD; PRT; 316 AA.  
 AC 059935;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE MANNOSE-6-PHOSPHATE ISOMERASE (EC 5.3.1.8) (PHOSPHOMANNOSE ISOMERASE)  
 DE (PMT) (PHOSPHOHEXOMUTASE).  
 GN PMT.  
 OS Streptococcus mutans.  
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 CC Streptococcus  
 CC NCBI\_TaxID=1309;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GS-5;  
 RX MEDLINE=94123951; PubMed=8293960;  
 RA Sato Y., Yamamoto Y., Kizaki H., Kuramitsu H.K.:  
 RT "Isolation and sequence analysis of the pm1 gene encoding  
 RT phosphomannose isomerase of Streptococcus mutans."  
 RL FEMS Microbiol. Lett. 114:61-66(1993).  
 CC -1- CATALYTIC ACTIVITY: D-MANNOSE 6-PHOSPHATE = D-FRUCTOSE  
 CC 6-PHOSPHATE.  
 CC -1- COFACTOR: ZINC (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF MANNOSE-6-PHOSPHATE ISOMERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: D16594; BAA04021.1; -  
 DR InterPro: IPR001250; -  
 DR Pfam: PF01238; PM1\_Type1; 1.  
 DR PROSITE: PS00965; PM1\_1; FALSE\_NEG.  
 DR PROSITE: PS00966; PM1\_1\_2; FALSE\_NEG.  
 KN Isomerase; Zinc.  
 FT METAL 97 97 ZINC (BY SIMILARITY).  
 FT METAL 109 109 ZINC (BY SIMILARITY).  
 FT METAL 171 171 ZINC (BY SIMILARITY).  
 SQ SEQUENCE 316 AA; 35388 MW; 8FA7B248791BB96A CRC64;

Query Match 63.8%; Score 37; DB 1; Length 316;  
 Best Local Similarity 55.6%; Pred. No. 46;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 KTMGOYWA 9  
 : || ||:  
 Db 32 ETTGEYWA 40

RESULT 19  
 CCSA\_PORPU STANDARD; PRT; 319 AA.  
 ID CCSA\_PORPU  
 AC P51369; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE CYTOCHROME C BIOGENESIS PROTEIN CCSA.  
 GN CCSA.  
 OS Porphyra purpurea.  
 CC Chloroplast.  
 CC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.  
 CC NCBI\_TaxID=2787;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AVONPORT;  
 RA Reith M.E., Munholland J.:  
 RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast  
 RT genome."  
 RL Plant Mol. Biol. Rep. 13:333-335(1995).  
 CC -1- FUNCTION: REQUIRED DURING CYTOCHROME BIOGENESIS AT THE STEP OF  
 CC HEME ATTACHMENT (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE CCMF/CYCK/CCIL/NRFE/CCSA FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

Query Match 63.8%; Score 37; DB 1; Length 319;  
 Best Local Similarity 50.0%; Pred. No. 47;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTMGOYWA 8  
 : || ||:  
 Db 248 EAMGSYMS 255

RESULT 20  
 CCSA\_PINTH STANDARD; PRT; 320 AA.  
 ID CCSA\_PINTH  
 AC P41650;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE CYTOCHROME C BIOGENESIS PROTEIN CCSA.  
 GN CCSA.  
 OS Pinus thunbergii (Green pine) (Japanese black pine).  
 CC Chloroplast.  
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Coniferopsida; Coniferales; Pinaceae; Pinus.  
 CC NCBI\_TaxID=3350;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95024047; PubMed=7937893;  
 RA Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,

RA Suglura M.;  
RT "Loss of all ndh genes as determined by sequencing the entire  
RT chloroplast genome of the black pine *Pinus thunbergii*.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).  
CC -1- FUNCTION: REQUIRED DURING CYTOCHROME BIOGENESIS AT THE STEP OF  
CC HEME ATTACHMENT (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE CCMF/CYCK/CCL1/NRFE/CCSA FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: D17510; BAA04448.1; -;  
CC InterPro: IPR002541; -;  
CC Pfam: PF01578; CytC\_asm; 1.  
CC Cytochrome c-type biogenesis; Chloroplast.  
KW SEQUENCE 320 AA; 36262 MW; 12D641127F4B306E CRC64;  
SQ  
Query Match 63.8%; Score 37; DB 1; Length 320;  
Best Local Similarity 50.0%; Pred. No. 47;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
OY 1 KTWGQYMA 8  
DB 248 EAWGSYS 255  
RESULT 21  
ID CCSA\_CYAPA STANDARD; PRT; 322 AA.  
AC P48257;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE CYTOCHROME C BIOGENESIS PROTEIN CCSA.  
GN CCSA.  
OS Cyanophora paradoxa.  
OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.  
OX NCBI\_Taxid=2762;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LB555 / PRINGSHEIM;  
RA Striwall V.L., Michalowski C.B., Luffelhardt W., Bohnert H.J.,  
RL Bryant D.A.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: REQUIRED DURING CYTOCHROME BIOGENESIS AT THE STEP OF  
CC HEME ATTACHMENT (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE CCMF/CYCK/CCL1/NRFE/CCSA FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U30821; AAA81298.1; -;  
CC InterPro: IPR002541; -;  
CC Pfam: PF01578; CytC\_asm; 1.  
CC Cytochrome c-type biogenesis; Cyanelle.  
KW SEQUENCE 322 AA; 37453 MW; 7E0CB2E5012A1838 CRC64;  
SQ  
Query Match 63.8%; Score 37; DB 1; Length 322;  
Best Local Similarity 50.0%; Pred. No. 47;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGQYMA 8  
DB 251 EAWGSYS 258  
RESULT 22  
ID CCSA\_CHLRE STANDARD; PRT; 353 AA.  
AC P48269;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE CYTOCHROME C BIOGENESIS PROTEIN CCSA.  
GN CCSA.  
OS Chlamydomonas reinhardtii.  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
CC Chlamydomonadaceae; Chlamydomonas.  
OX NCBI\_Taxid=3055;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CC-400;  
RX MEDLINE=95334509; PubMed=7610180;  
RA Chen Z.Y., Moroney J.V.;  
RT "Identification of a Chlamydomonas reinhardtii chloroplast gene with  
RT significant homology to bacterial genes involved in cytochrome c  
RT biosynthesis.";  
RL Plant Physiol. 108:843-844(1995).  
RN [2]  
RP SEQUENCE FROM N.A., AND FUNCTION.  
RC STRAIN=CC-503;  
RX MEDLINE=96214878; PubMed=8617725;  
RA Xie Z., Merchant S.;  
RT "The plastid-encoded *ccsa* gene is required for heme attachment to  
RT chloroplast c-type cytochromes.";  
RL J. Biol. Chem. 271:4632-4639(1996).  
CC -1- FUNCTION: REQUIRED DURING CYTOCHROME BIOGENESIS AT THE STEP OF  
CC HEME ATTACHMENT.  
CC -1- SIMILARITY: BELONGS TO THE CCMF/CYCK/CCL1/NRFE/CCSA FAMILY.  
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-20 IS THE INITIATOR.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U15556; AAA76600.1; -;  
CC InterPro: IPR002541; -;  
CC Pfam: PF01578; CytC\_asm; 1.  
CC Cytochrome c-type biogenesis; Chloroplast.  
KW COMPACT 342 342 T->I (IN REF. 2).  
FT SEQUENCE 353 AA; 39971 MW; 3D46ABC578044CB CRC64;  
SQ  
Query Match 63.8%; Score 37; DB 1; Length 353;  
Best Local Similarity 50.0%; Pred. No. 51;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
OY 1 KTWGQYMA 8  
DB 282 EAWGSYS 289  
RESULT 23  
ID SMF\_HAEIN STANDARD; PRT; 373 AA.  
AC P43862;  
DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (rel. 32, last sequence update)  
 DT 15-DEC-1998 (rel. 37, last annotation update)  
 DE SMF PROTEIN (DNA PROCESSING CHAIN A).  
 GN SMF OR DPRA OR H10985.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;  
 OC Haemophilus.  
 NCBI\_TaxID=727;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RD / KM20 / ATCC 51907;  
 RA Karandapuram S., Zhao X., Barcak G.J.;  
 RL Submitted (Dec-1994) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RD / KM20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 Kerlavage A.R., Sutton G., Fritch J.W., Fields C.A., Gocayne J.D.,  
 McKenney K., Smit G., Stacey G., Stacey G., Stacey G., Stacey G.,  
 Scott J.D., Shiley R., Liu L.-I., Glodex A., Kelley J.M.,  
 Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 Fine L.D., Fritchman J.L., Fuhrman J.L., Geoghegan N.S.M.,  
 Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 Venter J.C.;  
 RA "Whole-genome random sequencing and assembly of Haemophilus  
 influenzae Rd.";  
 RT Science 269:496-512(1995).  
 CC -1- FUNCTION: REQUIRED FOR CHROMOSOMAL BUT NOT PLASMID TRANSFORMATION.  
 CC -1- SIMILARITY: BELONGS TO THE SMF FAMILY.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL: U18657; AAA7011.1; -  
 DR EMBL: U32779; AAC22646.1; -  
 DR TIGR: H10985; -  
 SO SEQUENCE 373 AA; 41633 MW; 26E0353572EF61F CRC64;  
  
 Query Match 63.8%; Score 37; DB 1; Length 373;  
 Best Local Similarity 71.4%; Pred. No. 54;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 2 TWGQYWA 8  
 Db 123 TYGEYWA 129  
  
 RESULT 24  
 YG10\_ECOLI STANDARD; PRT; 739 AA.  
 AC 046861; Q46862;  
 DT 01-NOV-1997 (rel. 35, Created)  
 DT 01-OCT-2000 (rel. 40, Last sequence update)  
 DT 01-OCT-2000 (rel. 40, Last annotation update)  
 DE HYPOTHETICAL 83.5 KDA PROTEIN IN METC-SUPI INTERGENIC REGION.  
 GN YG10 OR B3015/B3016.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 NCBI\_TaxID=562;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN (2)  
 RP CONCEPTUAL TRANSLATION.  
 RA Rudd K.E.;  
 RL Unpublished observations (JUN-1999).  
 CC -1- SIMILARITY: TO M. JANNASCHII MJ1155. ALSO SOME SIMILARITY TO FAMILY  
 CC UPF0004.  
 CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT WAS  
 CC INTRODUCED IN POSITION 319 TO PRODUCE THIS ORF.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL: U28377; AAA69183.1; ALT\_FRAME.  
 DR EMBL: U28377; AAA69184.1; ALT\_FRAME.  
 DR EMBL: AE000383; AAC76052.1; ALT\_FRAME.  
 DR EMBL: AE000383; AAC76051.1; ALT\_FRAME.  
 DR Ecocore: EG13019; YG10.  
 KW Hypothetical protein.  
 SO SEQUENCE 739 AA; 83503 MW; ECE565EC8915384A CRC64;  
  
 Query Match 63.8%; Score 37; DB 1; Length 739;  
 Best Local Similarity 71.4%; Pred. No. 1e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 2 TWGQYWA 8  
 Db 15 SMPQYWA 21  
  
 RESULT 25  
 XYND\_RUMFL STANDARD; PRT; 802 AA.  
 AC 053317;  
 DT 01-NOV-1997 (rel. 35, Created)  
 DT 01-NOV-1997 (rel. 35, Last sequence update)  
 DT 15-JUL-1999 (rel. 38, Last annotation update)  
 DE XYLANASE/BETA-GLUCANASE PRECURSOR [INCLUDES: ENDO-1,4-BETA-XYLANASE  
 DE (EC 3.2.1.8) (XYLANASE); (ENDO-BETA-1,3-1,4 GLUCANASE) (EC 3.2.1.73)  
 DE (1,3-1,4-BETA-D-GLUCAN 4-GLUCANOHYDROLASE) (LICHENASE)].  
 GN XYND.  
 OS Ruminooccus flavefaciens.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
 OC Ruminooccus.  
 NCBI\_TaxID=1265;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-17;  
 RX MEDLINE=93259938; PubMed=8491715;  
 RA Flint H.J., Martin J., McPherson C.A., Daniel A.S., Zhang J.-X.;  
 RT "A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-  
 RT glucanase domains, encoded by the xynD gene of Ruminooccus  
 RT flavefaciens.";  
 RL J. Bacteriol. 175:2943-2951(1993).  
 CC -1- FUNCTION: CONTAINS TWO CATALYTIC DOMAINS WITH XYLANASE AND ENDO-  
 CC BETA-1,3-1,4 GLUCANASE ACTIVITIES.  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC  
 CC LINKAGES IN XYLANS.  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-BETA-D-GLYCOSIDIC LINKAGES  
 CC IN BETA-D-GLUCANS CONTAINING 1,3- AND 1,4-BONDS.  
 CC -1- PATHWAY: XYLAN DEGRADATION.  
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULOSE FAMILY  
 CC 6 (FAMILY II OF GLYCOSYL HYDROLASES).



CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO FAMILY 16 OF  
CC GLYCOSYL HYDROLASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; S61204; AAB26620.1; -  
DR HSSP; P23904; LAJK.  
DR InterPro; IPR000757; -  
DR InterPro; IPR001137; -  
DR Pfam; PF02018; CBD\_6; 1.  
DR Pfam; PF00457; Glyco\_hydro\_11; 1.  
DR Pfam; PF00722; Glyco\_hydro\_16; 1.  
DR PRINTS; PR00737; GLHYDRLASE16.  
DR PRINTS; PR00911; GLHYDRLASE11.  
DR PROSITE; PS00776; GLYCOSYL\_HYDROL\_F11\_1; 1.  
DR PROSITE; PS00777; GLYCOSYL\_HYDROL\_F11\_2; 1.  
DR PROSITE; PS01034; GLYCOSYL\_HYDROL\_F16; 1.  
DR Kxlan degradation; Hydrolase; Glycosidase; Signal;  
KM Multifunctional enzyme.  
FT SIGNAL; 1 31  
FT CHAIN; 32 802  
FT DOMAIN; 32 244  
FT DOMAIN; 245 523  
FT DOMAIN; 524 555  
FT DOMAIN; 556 802  
FT ACT\_SITE; 124 124  
FT ACT\_SITE; 226 226  
FT ACT\_SITE; 684 684  
FT DOMAIN; 524 529  
FT DOMAIN; 532 543  
FT DOMAIN; 546 553  
SQ SEQUENCE 802 AA; 89091 MW; 2880A689647284AF CRC64;  
  
Query Match 63.88; Score 37; DB 1; Length 802;  
Best Local Similarity 66.7%; Pred. No. 1.le+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
OY 1 KTWGYMAY 9  
Db 172 KTFPQWMSV 180  
  
RESULT 26  
YF68\_METJA STANDARD; PRT; 1009 AA.  
ID YF68\_METJA  
AC 058863;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL PROTEIN MJ1468.  
GN MJ1468.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
OC Methanococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSM 2661;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sult C.J., Blake O., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kierulff G.G., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weissstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,

RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;  
RT "complete genome sequence of the methanogenic archaeon, Methanococcus  
RT jannaschii.";  
RL Science 273:1058-1073(1996).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -1- SIMILARITY: CONTAINS 5 PKD DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; U67588; AAB99478.1; -  
DR TIGR; MJ1468; -  
DR InterPro; IPR000601; -  
DR Pfam; PF00801; PKD; 3.  
DR PROSITE; PS50093; PKD; 5.  
KM Hypothetical protein; Transmembrane; Repeat.  
FT TRANSMEM 6 26  
FT TRANSMEM 985 1005  
FT DOMAIN; 213 247  
FT DOMAIN; 436 503  
FT DOMAIN; 724 806  
FT DOMAIN; 822 886  
FT DOMAIN; 925 962  
FT DOMAIN; 293 298  
SQ SEQUENCE 1009 AA; 115119 MW; 13E9B4933EAB7972 CRC64;  
  
Query Match 63.88; Score 37; DB 1; Length 1009;  
Best Local Similarity 80.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
OY 3 WGYMAY 7  
Db 322 WGYMAY 326  
  
RESULT 27  
PA21\_BUNMU STANDARD; PRT; 120 AA.  
ID PA21\_BUNMU  
AC P00617;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE PHOSPHOLIPASE A2, BETA BUNGAROTOXIN A1 CHAIN (EC 3.1.1.4)  
DE (PHOSPHATIDYLCHOLINE 2-ACYLHYDROLASE).  
OS Bungarus multicinctus (Many-banded krait).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;  
OC Elapidae; Bungarinae; Bungarus.  
OX NCBI\_TaxID=8616;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Venom;  
RX MEDLINE=78109400; PubMed=624701;  
RA Kondo K., Narita K., Lee C.-Y.;  
RT "Amino acid sequences of the two polypeptide chains in beta-1-  
RT bungarotoxin from the venom of Bungarus multicinctus.";  
RL J. Biochem. 83:101-115(1978).  
RN [2]  
RP SEQUENCE, AND REVISIONS TO 85-87 AND 109.  
RX MEDLINE=82239269; PubMed=7096304;  
RA Kondo K., Toda H., Narita K., Lee C.-Y.;  
RT "Amino acid sequence of beta 2-bungarotoxin from Bungarus  
RT multicinctus venom. The amino acid substitutions in the B chains.";  
RL J. Biochem. 91:1519-1530(1982).  
RN [3]  
RP CHARACTERIZATION OF PHOSPHOLIPASE A2 ACTIVITY.  
RX MEDLINE=79088714; PubMed=730754;

RA Kondo K., Toda H., Narita K.;  
 RT "Characterization of phospholipase A activity of beta1-bungarotoxin  
 from Bungarus multicinctus venom. II. Identification of the histidine  
 residue of beta1-bungarotoxin modified by p-bromophenacyl bromide.";  
 RL J. Biochem. 84:1301-1308(1978).  
 [4]  
 RN CHARACTERIZATION OF PRESYNAPTIC NEUROTOXINS.  
 RP MEDLINE-78043174; PubMed-303565;  
 RX Abe T., Alema S., Miledi R.;  
 RT "Isolation and characterization of presynaptically acting neurotoxins  
 from the venom of Bungarus snakes.";  
 RL Eur. J. Biochem. 80:1-12(1977).  
 CC -1- FUNCTION: INHIBITS NEUROMUSCULAR TRANSMISSION BY BLOCKING  
 CC ACETYLCHOLINE RELEASE FROM THE NERVE TERMINI. ACT PRESYNAPTICALLY.  
 CC -1- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE  
 CC 2-ACYL GROUPS IN 3-SN-PHOSPHOLIPIDS.  
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCEROL-  
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.  
 CC -1- SUBUNIT: DIMER OF DISSIMILAR CHAINS LINKED BY ONE OR MORE  
 CC DISULFIDE BONDS. THE A CHAINS HAVE PHOSPHOLIPASE A2 ACTIVITY AND  
 CC THE B CHAINS SHOW HOMOLOGU WITH THE BASIC PROTEASE INHIBITORS. THE  
 CC A1 CHAIN IS FOUND IN BETA-1 AND BETA-2 BUNGAROTOXINS.  
 CC -1- MISCELLANEOUS: LD(50) IS 0.019 MG/KG BY INTRAPERITONEAL INJECTION  
 CC IN BETA-1 BUNGAROTOXIN AND 0.028 MG/KG BY INTRAPERITONEAL INJECTION  
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.  
 CC PIR: A00757; PSKFAU.  
 DR HSSP: P15445; 1A3F.  
 DR InterPro: IPR001211; -;  
 DR Pfam: PF00068; Phoslip; 1.  
 DR PRINTS: PR00389; PHPLIPASEA2.  
 DR PROSITE: PS00118; PA2\_HIS; 1.  
 DR PROSITE: PS00119; PA2\_ASP; 1.  
 KW Hydrolyase; Lipid degradation; Calcium; Multigene family; Venom;  
 KN Presynaptic neurotoxin.  
 FT ACT\_SITE 48 48 BY SIMILARITY.  
 FT ACT\_SITE 94 94 BY SIMILARITY.  
 FT DISULFID 15 15 INTERCHAIN (WITH A B CHAIN) (PROBABLE).  
 FT DISULFID 27 119 BY SIMILARITY.  
 FT DISULFID 29 45 BY SIMILARITY.  
 FT DISULFID 44 100 BY SIMILARITY.  
 FT DISULFID 51 93 BY SIMILARITY.  
 FT DISULFID 61 86 BY SIMILARITY.  
 FT DISULFID 79 91 BY SIMILARITY.  
 FT CA\_BIND 49 49 BY SIMILARITY.  
 FT VARIANT 89 89 I -> V (IN 20%).  
 SQ SEQUENCE 120 AA; 13489 MW; C6G544CE62CA12C2 CRC64;

Query Match 62.1%; Score 36; DB 1; Length 120;  
 Best Local Similarity 83.3%; Pred. No. 27;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGOY 6  
 ID PA22\_BUNMU STANDARD; PRT; 120 AA.  
 AC P00619;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE PHOSPHOLIPASE A2, BETA BUNGAROTOXIN A3 CHAIN (EC 3.1.1.4)  
 DE (PHOSPHATIDYLCHOLINE 2-ACYLHYDROLASE).  
 OS Bungarus multicinctus (Many-banded Krait).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidodermata; Squamata; Scleroglossa; Serpentes; Colubroidea;  
 OC Elapidae; Bungarinae; Bungarus.  
 OX NCBI\_TaxID=8616;  
 RN [1]  
 RM SEQUENCE.

RC TISSUE-Venom;  
 RX MEDLINE-82239270; PubMed-7096305;  
 RA Kondo K., Toda H., Narita K., Lee C.-Y.;  
 RT "Amino acid sequences of three beta-bungarotoxins (beta 3-, beta 4-,  
 RT and beta 5-bungarotoxins) from Bungarus multicinctus venom. Amino  
 RT acid substitutions in the A chains.";  
 RL J. Biochem. 91:1531-1548(1982).  
 CC -1- FUNCTION: INHIBITS NEUROMUSCULAR TRANSMISSION BY BLOCKING  
 CC ACETYLCHOLINE RELEASE FROM THE NERVE TERMINI. ACT PRESYNAPTICALLY.  
 CC -1- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE  
 CC 2-ACYL GROUPS IN 3-SN-PHOSPHOLIPIDS.  
 CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCEROL-  
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.  
 CC -1- SUBUNIT: DIMER OF DISSIMILAR CHAINS LINKED BY ONE OR MORE  
 CC DISULFIDE BONDS. THE A CHAINS HAVE PHOSPHOLIPASE A2 ACTIVITY AND  
 CC THE B CHAINS SHOW HOMOLOGU WITH THE BASIC PROTEASE INHIBITORS.  
 CC -1- MISCELLANEOUS: THE A3 CHAIN IS FOUND IN BETA-5 BUNGAROTOXIN.  
 CC -1- MISCELLANEOUS: LD(50) IS 0.13 MG/KG BY INTRAPERITONEAL INJECTION.  
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.  
 CC PIR: A00758; PSKFAU.  
 DR HSSP: P15445; 1A3F.  
 DR InterPro: IPR001211; -;  
 DR Pfam: PF00068; Phoslip; 1.  
 DR PRINTS: PR00389; PHPLIPASEA2.  
 DR PROSITE: PS00118; PA2\_HIS; 1.  
 DR PROSITE: PS00119; PA2\_ASP; 1.  
 KW Hydrolyase; Lipid degradation; Calcium; Multigene family; Venom;  
 KN Presynaptic neurotoxin.  
 FT ACT\_SITE 48 48 BY SIMILARITY.  
 FT ACT\_SITE 94 94 BY SIMILARITY.  
 FT DISULFID 15 15 INTERCHAIN (WITH A B CHAIN) (PROBABLE).  
 FT DISULFID 27 119 BY SIMILARITY.  
 FT DISULFID 29 45 BY SIMILARITY.  
 FT DISULFID 44 100 BY SIMILARITY.  
 FT DISULFID 51 93 BY SIMILARITY.  
 FT DISULFID 61 86 BY SIMILARITY.  
 FT DISULFID 79 91 BY SIMILARITY.  
 FT CA\_BIND 49 49 BY SIMILARITY.  
 SQ SEQUENCE 120 AA; 13439 MW; 0301BF045AB31770 CRC64;

Query Match 62.1%; Score 36; DB 1; Length 120;  
 Best Local Similarity 83.3%; Pred. No. 27;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGOY 6  
 ID PA22\_BUNMU STANDARD; PRT; 145 AA.  
 AC P00618;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE PHOSPHOLIPASE A2, BETA BUNGAROTOXIN A2 CHAIN PRECURSOR (EC 3.1.1.4)  
 DE (PHOSPHATIDYLCHOLINE 2-ACYLHYDROLASE).  
 OS Bungarus multicinctus (Many-banded Krait).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidodermata; Squamata; Scleroglossa; Serpentes; Colubroidea;  
 OC Elapidae; Bungarinae; Bungarus.  
 OX NCBI\_TaxID=8616;  
 RN [1]  
 RM SEQUENCE FROM N.A.  
 RP TISSUE-Venom gland;  
 RC MEDLINE-90356416; PubMed-2388842;  
 RX Danse J.M., Toussaint J.L., Kempf J.;  
 RT "Nucleotide sequence encoding beta-bungarotoxin A2-chain from the  
 RT venom glands of Bungarus multicinctus.";  
 RL Nucleic Acids Res. 18:4609-4609(1990).  
 RN [2]

```

RP      SEQUENCE OF 26-145.
RC      TISSUE=Venom:
RX      MEDLINE=82239270; PubMed=7096305;
RA      Kondo K., Toda H., Narita K., Lee C.-Y.;
RT      "Amino acid sequences of three beta-bungarotoxins (beta 3-, beta 4-,
RT      and beta 5-bungarotoxins) from Bungarus multicinctus venom. Amino
RT      acid substitutions in the A chains.";
RL      J. Biochem. 91:1531-1548(1982).
CC      -1- FUNCTION. INHIBITS NEURONIC TRANSMISSION BY BLOCKING
CC      ACETYLCHOLINE RELEASE FROM THE NERVE TERMINI. ACT PRESYNAPTICALLY.
CC      -1- FUNCTION. PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
CC      2-ACYL GROUPS IN 3-SN-PHOSPHOGlycerIDES.
CC      -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC      PHOSPHOCHOLINE + A FATTY ACID ANION.
CC      -1- SUBUNIT: DIMER OF DISSIMILAR CHAINS LINKED BY ONE OR MORE
CC      DISULFIDE BONDS. THE A CHAINS HAVE PHOSPHOLIPASE A2 ACTIVITY AND
CC      THE B CHAINS SHOW HOMOLOG WITH THE BASIC PROTEASE INHIBITORS. THE
CC      A2 CHAIN IS FOUND IN BETA-3 AND BETA-4 BUNGAROTOXINS.
CC      -1- MISCELLANEOUS: LD(50) IS 0.066 MG/KG BY INTRAPERITONEAL INJECTION
CC      IN BETA-3 BUNGAROTOXIN AND 0.073 MG/KG IN BETA-4.
CC      -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
CC      EMBL; X53407; CAAB37483.1; -.
DR      PIR; S10980; PSKFA2.
DR      HSSP; P15445; 1A3F.
DR      InterPro; IPR001211; -.
DR      Pfam; PF00068; phoslip; 1.
DR      PRINTS; PR00389; PHPHLIPASEA2.
DR      PROSITE; PS00118; PA2_HIS; 1.
DR      PROSITE; PS00119; PA2_ASP; 1.
KW      Hydrolyase; Lipid degradation; Calcium; Multigene family; Venom;
KW      Presynaptic neurotoxin; Signal.
FT      SIGNAL 1 25
FT      CHAIN 26 145
FT      PHOSPHOLIPASE A2, BETA BUNGAROTOXIN A2
FT      CHAIN.
FT      ACT_SITE 73 73 BY SIMILARITY.
FT      ACT_SITE 119 119 BY SIMILARITY.
FT      DISULFID 40 40 INTERCHAIN (WITH A B CHAIN) (PROBABLE).
FT      DISULFID 52 144 BY SIMILARITY.
FT      DISULFID 54 70 BY SIMILARITY.
FT      DISULFID 69 125 BY SIMILARITY.
FT      DISULFID 76 118 BY SIMILARITY.
FT      DISULFID 86 111 BY SIMILARITY.
FT      DISULFID 104 116 BY SIMILARITY.
FT      CA_BIND 74 74 BY SIMILARITY.
FT      CONFLICT 91 92 OS -> SO (IN REF. 2).
FT      CONFLICT 128 128 N -> Q (IN REF. 2).
FT      CONFLICT 130 130 E -> D (IN REF. 2).
SO      SEQUENCE 145 AA; 16296 MW; 08CD9D0E4A57581 CRC64;

```

Query Match	62.1%	Score 36	DB 1	Length 145
Best Local Similarity	83.3%	Pred. No. 32		
Matches	5	Conservative	1	Mismatches 0
				Gaps 0

Db 42 KTWGEY 47

RESULT	30	
PA24_BUNMU		
ID	PA24_BUNMU	STANDARD;
AC.	P17934;	PRT; 147 AA.
DT	01-NOV-1990 (Rel. 16, Created)	

DT	01-NOV-1990 (Rel. 16, Last sequence update)
DE	15-JUL-1999 (Rel. 38, Last annotation update)
OS	PHOSPHOLIPASE A2, BETA BUNGAROTOXIN A4 CHAIN PRECURSOR (EC 3.1.1.4)
DE	(PHOSPHATIDYLCHOLINE 2-ACYLHYDROLASE).
OS	Bungarus multicinctus (Many-banded krait).
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Lepidosauria; Squamata; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Elapidae; Bungarinae; Bungarus.
CC	NCBI_TaxID=8616;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Venom gland;
RX	MEDLINE=90356417; PubMed=2388843;
RA	Danse J.M., Garnier J.M., Kempf J.;
RT	"cDNA deduced amino-acid sequence of a new phospholipase from
RL	Bungarus multicinctus";
RL	Nucleic Acids Res. 18:4610-4610(1990).
CC	-1- FUNCTION: INHIBITS NEUROMUSCULAR TRANSMISSION BY BLOCKING
CC	ACETYLCHOLINE RELEASE FROM THE NERVE TERMINI. ACT PRESYNAPTICALLY.
CC	-1- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
CC	2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
CC	-1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC	PHOSPHOCHOLINE + A FATTY ACID ANION.
CC	-1- SUBUNIT: DIMER OF DISSIMILAR CHAINS LINKED BY ONE OR MORE
CC	DISULFIDE BONDS. THE A CHAINS HAVE PHOSPHOLIPASE A2 ACTIVITY AND
CC	THE B CHAINS SHOW HOMOLOGY WITH THE BASIC PREPARE INHIBITORS.
CC	-1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; X53408; CAA37484.1; -
DR	PIR; S10982; PSKFA4.
DR	HSSP; P15445; 1A3F.
DR	InterPro: IPR001211; -
DR	Pfam; PF00068; phoslip.1.
DR	PRINTS; PR00389; PHPLIPASEA2.
DR	PROSITE; PS00118; PA2_HIS; 1.
DR	PROSITE; PS00119; PA2_ASP; 1.
KW	Hydrolase; Lipid degradation; Calcium; Multigene family; Venom;
KW	Presynaptic neurotoxin; Signal.
FT	SIGNAL
FT	1 27
FT	CHAIN
FT	28 147
FT	PHOSPHOLIPASE A2, BETA BUNGAROTOXIN A4
FT	CHAIN.
FT	BY SIMILARITY.
FT	ACT SITE 75 75
FT	ACT SITE 119 119
FT	DISULFID 42 42
FT	DISULFID 54 146
FT	DISULFID 56 72
FT	DISULFID 71 127
FT	DISULFID 78 120
FT	DISULFID 88 113
FT	DISULFID 106 118
FT	CA_BIND 76 76
FT	SEQUENCE 147 AA; 16177 MW; 200CAF65F8CE2B6 CRC64;

Query Match	62.1%	Score 36	DB 1	Length 147
Similarity	83.3%	Pred No. 33		
Best Local				
Matches	5	Conservative	1	Mismatches 0
				Indels 0
				Gaps 0

QY	1	KTWGQY	6
		:	1
Db	44	KTWGEX	49

Search completed: July 3, 2001, 10:28:37

Tue Jul 3 20:10:08 2001

us-09-214-836-1.rsp

Page 16

Job time: 322 sec

7

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:00:55 ; Search time 33.34 Seconds  
(without alignments)  
16.365 Million cell updates/sec

Title: US-09-214-836-1

Perfect score: 58

Sequence: 1 KTWGQYNAV 9.

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_0601:\*

- 1: /SIDSR/gcgdata/geneseq/AA1980.DAT:\*
- 2: /SIDSR/gcgdata/geneseq/AA1981.DAT:\*
- 3: /SIDSR/gcgdata/geneseq/AA1982.DAT:\*
- 4: /SIDSR/gcgdata/geneseq/AA1983.DAT:\*
- 5: /SIDSR/gcgdata/geneseq/AA1984.DAT:\*
- 6: /SIDSR/gcgdata/geneseq/AA1985.DAT:\*
- 7: /SIDSR/gcgdata/geneseq/AA1986.DAT:\*
- 8: /SIDSR/gcgdata/geneseq/AA1987.DAT:\*
- 9: /SIDSR/gcgdata/geneseq/AA1988.DAT:\*
- 10: /SIDSR/gcgdata/geneseq/AA1989.DAT:\*
- 11: /SIDSR/gcgdata/geneseq/AA1990.DAT:\*
- 12: /SIDSR/gcgdata/geneseq/AA1991.DAT:\*
- 13: /SIDSR/gcgdata/geneseq/AA1992.DAT:\*
- 14: /SIDSR/gcgdata/geneseq/AA1993.DAT:\*
- 15: /SIDSR/gcgdata/geneseq/AA1994.DAT:\*
- 16: /SIDSR/gcgdata/geneseq/AA1995.DAT:\*
- 17: /SIDSR/gcgdata/geneseq/AA1996.DAT:\*
- 18: /SIDSR/gcgdata/geneseq/AA1997.DAT:\*
- 19: /SIDSR/gcgdata/geneseq/AA1998.DAT:\*
- 20: /SIDSR/gcgdata/geneseq/AA1999.DAT:\*
- 21: /SIDSR/gcgdata/geneseq/AA2000.DAT:\*
- 22: /SIDSR/gcgdata/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	100.0	9	19	AAW45771
2	53	91.4	9	16	AAW78644
3	53	91.4	9	19	AAW77119
4	53	91.4	9	19	AAW78850
5	53	91.4	9	19	AAW70010
6	53	91.4	9	19	AAW54598
7	53	91.4	9	19	AAW45777
8	53	91.4	9	19	AAW45770
9	53	91.4	9	20	AAW49663
10	53	91.4	9	20	AAW58524
11	53	91.4	9	20	AAW47616

12	53	91.4	9	20	AAW33172	Human gp100-Pmel11
13	53	91.4	9	20	AAW40211	Amino acid sequenc
14	53	91.4	9	20	AAW26867	Melanoma-derived I
15	53	91.4	9	20	AAW01753	Exemplary antigen I
16	53	91.4	9	20	AAW00715	Tumour antigen boe
17	53	91.4	9	20	AAW10449	HLA Class I motif
18	53	91.4	9	21	AAW33662	MHC class I associ
19	53	91.4	9	21	AAW23679	Cytotoxic T lymphi
20	53	91.4	9	21	AAW08694	Antigenic peptide
21	53	91.4	9	21	AAW71520	Human gp100 Pmel11
22	53	91.4	9	21	AAW02622	Tumour associated
23	53	91.4	9	21	AAW90803	Human leukocyte an
24	53	91.4	9	21	AAW92299	gp100-Pmel11 anti
25	53	91.4	9	21	AAW84296	Tumour associated
26	53	91.4	9	21	AAW82879	gp100/Pmel117 tum
27	53	91.4	9	21	AAW56614	gp100-Pmel-117 gen
28	53	91.4	9	22	AAW31354	Exemplary antigen
29	53	91.4	10	16	AAW78643	Immunogenic peptid
30	53	91.4	10	16	AAW84209	gp100 melanoma ant
31	53	91.4	11	16	AAW78642	Immunogenic peptid
32	53	91.4	30	22	AAW61647	gp100 peptide #1.
33	53	91.4	661	16	AAW84855	MART-1 melanoma an
34	53	91.4	661	16	AAW78646	Melanoma associate
35	53	91.4	661	20	AAW31977	Human melanoma ant
36	53	91.4	662	20	AAW42627	Human melanoma ant
37	53	91.4	668	18	AAW38164	Pmel17 encoded by
38	52	89.7	9	19	AAW45775	Melanoma associate
39	52	89.7	9	19	AAW45776	Melanoma associate
40	50	86.2	9	19	AAW42542	Gp 100 epitope (re
41	50	86.2	9	20	AAW46537	Immunogenic peptid
42	50	86.2	9	21	AAW13727	Peptide fragment #
43	50	86.2	10	16	AAW84208	gp100 melanoma ant
44	50	86.2	661	16	AAW84854	MART-1 melanoma an
45	49	84.5	625	20	AAW31978	Mouse melanoma ant

#### ALIGNMENTS

RESULT 1	
AAW45771	AAW45771 standard; peptide; 9 AA.
XX	XX
AC	AAW45771;
XX	XX
DT	22-JUN-1998 (first entry)
XX	XX
DE	Melanoma associated peptide analogue #2.
XX	XX
KW	Metastatic melanoma; peptide analogue; vaccine; cancer; diagnosis;
XX	XX
OS	Synthetic.
OS	Homo sapiens.
XX	XX
PN	WC09802538-A1.
XX	XX
PD	22-JAN-1998.
XX	XX
PF	08-JUL-1997; 97WO-EP03712.
XX	XX
PR	11-JUL-1996; 96EP-0201945.
XX	XX
PA	(ALKU) AK20 NOBEL NV.
XX	XX
PI	Adema GJ, Figdor CG;
XX	XX
DR	WPI; 1998-110586/10.
XX	XX
PT	Melanoma associated peptide analogues - useful in vaccines against
PT	melanoma
XX	XX
PS	Claim 4; Figure 1; 47pp; English.

XX This sequence represents a specifically claimed example of a novel  
CC peptide, which is immunogenic with lymphocytes directed against  
CC metastatic melanomas. It is characterized in that it comprises at least  
CC a part of the following sequence, where the amino acid at position 2 or 8  
CC is substituted: Lys-Thr-Tyr-Gln-Tyr-Trp-Gln-Val. Vaccines comprising  
CC the peptide, an epitope of the peptide, nucleotide sequence encoding the  
CC peptide, or an antigen presenting cell preloaded with the peptide or  
CC antibody as above, are useful for cancer, particularly melanoma,  
CC treatment. The peptides can also be used to generate antigen reactive  
CC tumour infiltrating lymphocytes, which can also be used in vaccines. The  
CC peptides can be exploited to elicit native epitope-reactive CTL. Usage  
CC of the peptides with improved immunogenicity may contribute to the  
CC development of CTL-epitope based vaccines in viral disease and cancer.  
CC  
SQ Sequence 9 AA;

Query Match 100.0%; Score 58; DB 19; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYNAV 9  
| | | | | | | | |  
DB 1 ktwgqywav 9

RESULT 2  
AAW78644  
ID AAW78644 standard; Protein; 9 AA.  
XX  
AC AAW78644;  
XX  
DT 22-JAN-1996 (first entry)  
XX  
XX Immunogenic peptide of melanoma associated antigen gp100.  
DE  
XX Melanoma; antigen; vaccine; immunogen; primer; probe; detection;  
KM Identification; tumour; gp100.  
XX  
XX Homo sapiens.  
OS  
XX  
XX EP668350-A1.  
PN  
XX  
XX 23-AUG-1995.  
PD  
XX  
XX 14-FEB-1995; 95EP-0200348.  
PF  
XX  
XX 21-DEC-1994; 94EP-0203709.  
PR  
XX 16-FEB-1994; 94EP-0200337.  
PR  
XX  
XX (ALKU ) AKZO NOBEL NV.  
PA  
XX  
XX Adema CJ, Figdor CG;  
PI  
XX  
XX WPI; 1995-284790/38.  
DR  
XX N-PSDB; AAQ96055.  
DR  
XX  
XX Melanoma associated antigen gp100 - used in vaccines and for the  
PT detection of tumours  
PT  
XX  
XX Claim 5; Page 31; 40pp; English.  
PS  
XX  
XX Immunogenic peptides derived from the melanoma associated antigen  
CC (See AAW78639-45) may be used in the production of vaccines.  
CC Nucleotide sequences encoding the immunogenic peptides may be used  
CC as primers and probes in the detection of melanoma cells. Tumour  
CC infiltrating lymphocytes capable of binding to the melanoma  
CC associated antigen can be cultured ex vivo and returned to melanoma  
CC particles, and when radiolabelled, they may be used to identify  
CC tumour deposits.  
CC  
SQ Sequence 9 AA;

Query Match 91.4%; Score 53; DB 16; Length 9;  
Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYNAV 9  
| | | | | | | | |  
DB 1 ktwgqywav 9

RESULT 3  
AAW77119  
ID AAW77119 standard; peptide; 9 AA.  
XX  
AC AAW77119;  
XX  
XX 16-NOV-1998 (first entry)  
DT  
XX  
XX gp100/Pmel17 synthetic peptide epitope 1.  
DE  
XX  
XX Tyrosinase; tyrosinase cytotoxic lymphocyte response;  
KM cytotoxic T lymphocyte; cysteine-depleted; melanoma.  
KW  
XX  
XX Synthetic.  
OS  
XX  
XX WO9833810-A2.  
PN  
XX  
XX 06-AUG-1998.  
PD  
XX  
XX 29-JAN-1998; 98WO-US01592.  
PF  
XX  
XX 30-JAN-1997; 97US-0037781.  
PR  
XX  
XX (UYVI-) UNIV VIRGINIA PATENT FOUND.  
PA  
XX  
XX Engelhard VH, Hunt DF, Kittlesen D, Slingluff CL;  
PI  
XX  
XX WPI; 1998-437388/37.  
DR  
XX  
XX Disease specific immunogen - comprises disease specific cytotoxic T  
PT lymphocyte epitope used to elicit melanoma specific CTL response  
PT  
XX  
XX Disclosure; Page 27; 93pp; English.  
PS  
XX  
XX The peptide epitope AAW77119-W77138 were created for human  
CC tumour-specific cytotoxic T lymphocyte response. These peptides are  
CC cysteine-depleted mutants of a native disease-specific CTL epitope. The  
CC response than the native epitope. The epitopes can be used in a  
CC disease-specific immunogen to protect a mammal against disease in  
CC particular melanomas. The peptides may also be used to screen a sample  
CC for the presence of an antigen with the same epitope, or with a different  
CC cross-reactive epitope.  
CC  
SQ Sequence 9 AA;

Query Match 91.4%; Score 53; DB 19; Length 9;  
Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYNAV 9  
| | | | | | | | |  
DB 1 ktwgqywav 9

RESULT 4  
AAW78850  
ID AAW78850 standard; peptide; 9 AA.  
XX  
AC AAW78850;  
XX

DT	17-NOV-1998	(first entry)
DE	PMEL 17 (GP100) protein fragment 154-162.	
XX		
XX		
XX		
KM	Microparticle; delivery; polymeric matrix; autoantigen; tumour antigen;	
KW	class II associated peptide; pathogen; gene therapy; genetic disease;	
XX	infection; downregulation; immune response.	
XX		
OS	Homo sapiens.	
XX	Synthetic.	
PN	WO9831398-A1.	
XX		
PD	23-JUL-1998.	
XX		
PF	22-JAN-1998; 98WO-US01499.	
XX		
PR	06-JAN-1998; 98US-0003253.	
PR	22-JAN-1997; 97US-0787547.	
PA	(PANG-) PANGAEA PHARM INC.	
XX		
PI	Curley JM, Hedley ML, Langer RS, Lunsford LB;	
DR	WPI; 1998-427556/36.	
XX		
PT	New preparations of microparticles - comprising a synthetic polymer	
PT	matrix and nucleic acid comprising an expression vector for use in	
XX	gene therapy	
PS	Disclosure; Page 10; 101pp; English.	
XX		
CC	A microparticle preparation (MP) has been developed, consisting of	
CC	microparticles having a diameter of less than 100 nm. The MP	
CC	comprises: (a) a polymeric matrix (PM) consisting of one or more	
CC	synthetic polymers having a solubility in water of less than 1 mg/l; and	
CC	(b) an expression vector selected from RNA molecules (at least 50% of	
CC	which are closed circles) or circular plasmid DNA (at least 50% of which	
CC	are supercoiled). Also described is a MP of at most 20 microns in	
CC	diameter, comprising: (a) a PM; and (b) a NM comprising an expression	
CC	control sequence operatively linked to a coding sequence, where the	
CC	coding sequence encodes an expression product selected from: (1) a	
CC	polypeptide at least 7 amino acids in length, having a sequence identical	
CC	to the sequence of: (1) a fragment of a naturally-occurring mammalian	
CC	protein; or (11) a fragment of a naturally-occurring protein from an	
CC	infectious agent which infects a mammal; (2) a peptide having a length	
CC	and sequence which permits it to bind to an MHC class I or II molecule;	
CC	and (3) the polypeptide or the peptide linked to a trafficking sequence.	
CC	AAW69763 to AAW69765, and AAW78793 to AAW78897 are peptide fragments for	
CC	use in the present invention. The MPs are highly effective vehicles for	
CC	the delivery of polynucleotides into phagocytic cells. They can be used	
CC	for gene therapy, e.g. for treating genetic diseases, infections or	
CC	tumours or for downregulating an immune response.	
XX		
SO	Sequence 9 AA:	
OY	Query Match 91.4%; Score 53; DB 19; Length 9;	
	Best Local Similarity 88.9%; Pred. No. 3.4e+05;	
DB	Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0	
	1 KTWGQYWAY 9	
	1 KTWGQYWAY 9	
RESULT 5		
AAW70010		
ID	AAW70010 standard; peptide; 9 AA.	
AC	AAW70010;	
XX		
DT	22-OCT-1998 (first entry)	

XX	Melanoma-associated antigen gp100 derived HLA-A2.1 binding peptide 1.
XX	Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;
KW	human leukocyte antigen; HLA; tumour associated antigen; cancer;
KM	antigen presenting cell; APC; immunogenic peptide; immune disorder;
KV	viral infection; AIDS; hepatitis; bacterial infection; malaria;
KW	fungal infection; tuberculosis; melanoma; gp100.
XX	Synthetic.
OS	Homo sapiens.
XX	MO98J386B-AL.
PN	06-AUG-1998.
XX	
PD	30-JAN-1998; 98WO-USO1959.
XX	
PF	31-JAN-1997; 97US-0036696.
PR	(EPLM-) EPIMADNE INC.
XX	
PA	Cells E, Sette A, Sidney J, Southwood S, Tsai V;
XX	
PI	WPt_1998-437445/37.
DR	
XX	Production of antigen-specific cytotoxic T cells - by incubating
PT	immunogenetic peptide(s) from antigen that binds class I major
PT	histocompatibility complex molecules with pre-treated antigen
PT	presenting cells
PS	
Example 4:	Page 62; 104pp; English.
XX	
CC	Sequences shown in AAWT0010 to AAWT0026 represent peptides derived from
CC	melanoma-associated antigen gp100 that can bind to a human leukocyte
CC	antigen (HLA). HLA-A2.1. The peptides are used to exemplify the method
CC	of invention of producing antigen-specific cytotoxic T cells (CTLs) in
CC	vitro. The method comprises contacting immunogenic peptides from an
CC	antigen that binds class I major histocompatibility complex (MHC)
CC	molecules with antigen presenting cells (APCs) pretreated with
CC	pretreatment growth factors, and incubating the APCs with purified CD8
CC	cells in the presence of at least 2 incubation growth factors, thereby
CC	producing antigen-specific CTLs. A method for specifically killing
CC	target cells in a human patient is also provided which comprises
CC	obtaining a fluid sample containing CTLs from a patient, contacting the
CC	cytotoxic T cells with APCs pretreated with pre-treatment growth
CC	factors, where the APCs comprise class I MHC molecules. The pretreated
CC	APCs are incubated with the cytotoxic growth factors, thereby producing
CC	activated CTLs which are contacted with a carrier to form a composition.
CC	The composition can then be administered to the patient. The activated
CC	CTLs can be used for treating cancers, immune disorders, viral
CC	infections, AIDS, hepatitis, bacterial infection, fungal infection,
CC	malaria or tuberculosis.
CC	
Sequence	9 AA: .  . .  Query Match Best Local Similarity 91.4%; Score 53; DB 19; Length 9; Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0
Oy	1 KTWGQYNAV 9 
Db	1 ktwgqymv 9
RESULT	6
ID	AAMS4598 standard; peptide; 9 AA.
AC	AAMS4598;
DT	25-SEP-1998 (first entry)

XX		Peptide 3 from gp 100/Pmel-17.	
DE		Mannose; antigen; antigen-presenting cell; mannosylated peptide; T cell;	
KM		vaccine; treatment.	
KW		Synthetic.	
OS		WO9813378-A1.	
PN		02-APR-1998.	
PD		25-SEP-1997;	
XX		97WO-NL00536.	
PF		26-SEP-1996;	
PR		96EP-0202701.	
XX		(UYLE-) RIJKSUNIV LEIDEN.	
PA		Drijfhout JW, Konig F;	
XX		WPI; 1998-230631/20.	
DR		Increasing uptake and presentation of antigen(s) - by adding mannose	
XX		residue(s) to antigen for increasing T cell response, useful in,	
PT		e.g. vaccines against viral infection(s)	
XX		Disclosure; Page 24; 47pp; English.	
PS		The peptides AAW54559-W54809 are examples of peptides to which at least	
CC		1 (preferably 2) mannose can be attached to increase their uptake as	
CC		antigen by antigen-presenting cells. Uptake of agonist mannosylated	
CC		peptides will increase the T cell response, whereas uptake of antagonist	
CC		peptides blocks the T cell response. Blocking binding of immunogenic	
CC		autoantigens can be used in treatment of type I diabetes, rheumatoid	
CC		arthritis, graft rejection etc., also to induce T-cell non-	
CC		responsiveness. Vaccines containing mannosylated antigen are used to	
CC		prevent or treat infections by, e.g. bacteria, viruses, fungi, helminths	
CC		and parasites.	
SO		Sequence    9 AA;	
QY		Query Match                  91.4%; Score 53; DB 19; Length 9;	
		Best Local Similarity      88.9%; Pred. No. 3.4e+05;	
		Matches     8; Conservative       0; Mismatches     1; Indels       0; Gaps       0.	
DB		1 KTWGQYWAV 9           1 ktwgqyqv 9	
RESULT    7			
AAMW5777		AAW45777 standard; peptide; 9' AA.	
AC		AAW45777;	
DT		22-JUN-1998 (first entry)	
DE		Melanoma associated peptide analogue #8.	
KV		Metastatic melanoma; peptide analogue; vaccine; cancer; diagnosis;	
KW		antigen; CTL; immunogenic; viral disease.	
OS		Synthetic.	
OS		Homo sapiens.	
XX		WO9802538-A1.	
PN		22-JAN-1998.	
PD		08-JUL-1997;    97WO-EP03712.	
XX			
PF			
XX			

```

PR    11-JUL-1996;      96EP-0201945.
XX
PA    (ALKU ) AKZO NOBEL NV.
XX
PI    Adema GJ, Figdor CG;
XX
DR    WPI; 1998-110586/10.
XX
PT    Melanoma associated peptide analogues - useful in vaccines against
XX melanoma
XX
PS    Claim 4; Page 35; 47pp; English.
CC
CC    This sequence represents a specifically claimed example of a novel
CC peptide which is immunogenic with lymphocytes directed against
CC metastatic melanomas. It is characterised in that it comprises at least
CC a part of the following sequence, where the amino acid at position 2 or 8
CC is substituted: Lys-Thr-Tyr-Gly-Gln-Tyr-Tyr-Gln-Val. Vaccines comprising
CC the peptide, an epitope of the peptide, nucleotide sequence encoding the
CC peptide, or an antigen presenting cell preloaded with the peptide or
CC antibody as above, are useful for cancer, particularly melanoma.
CC treatment. The peptides can also be used to generate antigen reactive
CC tumour infiltrating lymphocytes, which can also be used in vaccines. The
CC peptides can be exploited to elicit native epitope-reactive CTL. Usage
CC of the peptides with improved immunogenicity may contribute to the
CC development of CTL-epitope based vaccines in viral disease and cancer.
XX
SQ    Sequence      9 AA:
OY
OY    1 KTWGQYWAY 9
     | |||||
Db    1 kwygqyway 9
RESULT      8
AAM45770
ID AAM45770 standard; peptide; 9 AA.
AC AAM45770;
XX
XX    22-JUN-1998 (first entry)
DT
XX
XX    Melanoma associated peptide analogue #1.
DE
XX
XX    Metastatic melanoma; peptide analogue; vaccine; cancer; diagnosis;
KW antigen; CTL; Immunogenic; viral disease.
OS Synthetic.
OS Homo sapiens.
XX
PN WO9802538-A1.
XX
PD 22-JAN-1998.
XX
XX    08-JUL-1997; 97WO-EP03712.
PF
XX    11-JUL-1996; 96EP-0201945.
FR
XX    (ALKU ) AKZO NOBEL NV.
PA
PI Adema GJ, Figdor CG;
XX
XX    WPI; 1998-110586/10.
DR
XX
XX    Melanoma associated peptide analogues - useful in vaccines against
PT melanoma
XX
PS    Claim 1; Figure 1; 47pp; English.

```



XX This sequence represents a specifically claimed example of a novel  
CC peptide, which is immunogenic with lymphocytes directed against  
CC metastatic melanomas. It is characterised in that it comprises at least  
CC a part of the following sequence, where the amino acid at position 2 or 8  
CC is substituted: Lys-Thr-Trp-Gln-Tyr-Trp-Gln-Val. Vaccines comprising  
CC the peptide, an epitope of the peptide, nucleotide sequence encoding the  
CC peptide, or an antigen presenting cell preloaded with the peptide or  
CC antibody as above, are useful for cancer, particularly melanoma,  
CC treatment. The peptides can also be used to generate antigen reactive  
CC tumour infiltrating lymphocytes, which can also be used in vaccines. The  
CC peptides can be exploited to elicit native epitope-reactive CTL. Usage  
CC of the peptides with improved immunogenicity may contribute to the  
CC development of CTL-epitope based vaccines in viral disease and cancer.  
XX

SQ Sequence 9 AA:

Query Match 91.4%; Score 53; DB 19; Length 9;  
Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYMAV 9  
Db 1 ktwgqywmv 9  
|||||||  
1

RESULT 9  
AAV49663  
ID AAV49663 standard; peptide; 9 AA.  
XX  
AC AAV49663:  
XX  
DT 14-JAN-2000 (first entry)  
XX  
DE Tumour antigenic peptide SEQ ID NO:30.  
XX  
KW Human; sdph3.10; SAGE; sdg3.8; HAGE; sdg3.5; TRAP; sarcoma;  
KM tumour rejection antigen precursor; tumour associated nucleic acid;  
KW carcinoma; cancer; immune response; diagnosis.  
XX  
OS Synthetic.  
XX  
PN WO9953061-A2.  
XX  
PD 21-OCT-1999.  
XX  
PF 14-APR-1999; 99WO-US08163.  
XX  
PR 15-APR-1998; 98US-0060706.  
PR 27-JUL-1998; 98US-0122989.  
PR 30-OCT-1998; 98US-0183706.  
PR 30-OCT-1998; 98US-0183789.  
XX  
PA (LUDWIG) INST CANCER RES.  
XX  
PI Martelange V, De Smet C, Boon-Falleur T;  
XX  
DR WPI; 1999-620430/53.  
XX  
PT New nucleic acid encoding sarcoma-associated gene products, useful for  
PT diagnosing, e.g. treating and preventing cancer -  
XX  
PS Disclosure; Page 25; 93pp; English.

XX The present invention describes sarcoma-associated gene products (I).  
CC Agents, specifically sarcoma associated nucleic acids (II) or their  
CC expression products that are tumour rejection antigens (TRA), that  
CC selectively increase formation of HLA (human leucocyte antigen)/(I)  
CC complexes are used for treating cancer, especially sarcoma and  
CC carcinoma, in humans and other animals. Compositions containing  
CC autologous cytolytic T cells (CTL), specific for the HLA/(I) complex,  
CC are similarly useful, also transformed cells that stimulate such CTL

CC in vivo. (II) are also used: (i) as source of therapeutic antisense  
CC sequences that reduce expression of (II); (ii) for recombinant  
CC production of (I); (iii) particularly its fragments, as primers and  
CC probes in usual hybridisation and amplification assays, for diagnosis,  
CC prognosis and monitoring of tumours, or for measuring binding  
CC specificity of HLA molecules or CTL clones; (iv) to identify related  
CC sequences; and (v) for generating transgenic animals, e.g. for studying  
CC cancer and immune responses to it. (I) are used to raise specific  
CC antibodies (Ab) and therapeutically, Ab are used to diagnose tumours in  
CC immunoassays, also for delivering drugs, toxins, imaging agents etc. to  
CC (I)-expressing cells. AAV49637 to AAV49670 represent exemplary tumour  
CC antigenic peptides given in the present invention.  
XX

SQ Sequence 9 AA:

Query Match 91.4%; Score 53; DB 20; Length 9;  
Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYMAV 9  
Db 1 ktwgqywmv 9  
|||||||  
1

RESULT 10  
AAV53524  
ID AAV53524 standard; Protein; 9 AA.  
XX  
AC AAV53524:  
XX  
DT 18-JAN-2000 (first entry)  
XX  
DE Human melanoma Pmel17 (gp100) (aa 154-162) binds HLA-A2.  
XX  
KW Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;  
KM electrical charge; hydrophilicity; Vaccine; immune response; HIV; HBV;  
KW human immunodeficiency virus; hepatitis B virus; papilloma virus;  
KW melanoma; malaria; parasite.  
XX  
OS Synthetic.  
XX  
OS Homo sapiens.  
XX  
PN FR2776926-A1.  
XX  
PD 08-OCT-1999.  
XX  
PF 07-APR-1998; 98FR-0004323.  
XX  
PR 07-APR-1998; 98FR-0004323.  
XX  
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.  
PA (CNRS) CNRS CENT NAT RECH SCI.  
PA (INSP) INST PASTEUR LILLE.  
XX  
PI Le Gal FA, Guillet JG, Gahery SH, Gras MH, Melnyk O, Tartar A;  
XX  
DR WPI; 1999-583113/50.  
XX  
PT New lipopeptide containing lipid regions and two epitopes, all  
PT separated by peptide spacers that impart hydrophilicity, useful in  
PT vaccines -  
XX  
PS Disclosure; Page 24; 35pp; French.

XX The invention relates to the generation of a lipopeptide comprising at  
CC least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)  
CC epitope and at least one lipid residue with (i) the epitopes and lipid  
CC portion and (ii) the epitopes, being separated independently by peptide  
CC spacers. These spacers comprise sequences of amino acids which carry an  
CC overall electrical charge in neutral media to ensure that the  
CC lipopeptide is hydrophilic. The peptides AAV53301-Y53549 represents  
CC examples of peptide epitopes used to generate the lipopeptides. These are

CC used in therapeutic or prophylactic compositions and vaccines to induce  
CC specific immune responses against human immunodeficiency, hepatitis B or  
CC papilloma viruses; p53 of melanoma or the malaria parasite.

XX  
SQ Sequence 9 AA;

Query Match 91.4%; Score 53; DB 20; Length 9;  
Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYMAV 9  
111111111  
Db 1 KTWGQYMAV 9

## RESULT 11

AAV47616  
ID AAV47616 standard; Peptide: 9 AA.

XX  
AC AAV47616;

XX  
DT 01-DEC-1999 (first entry)

XX  
DE Immunogenic peptide having a human leukocyte antigen binding motif #2227.

XX  
KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;  
KW immune response; T cell activation; major histocompatibility complex;  
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;  
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;  
KW vaccine; immunisation.

XX  
OS Synthetic.

XX  
OS Homo sapiens.

XX  
PN WO945354-A1.

XX  
PD 16-SEP-1999.

XX  
PF 13-MAR-1998; 98WO-US05039.

XX  
PR 13-MAR-1998; 98WO-US05039.

XX  
PA (EPTM-) EPTIMUNE INC.

XX  
PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;

XX  
DR WPI: 1999-551214/46.

XX  
PT New immunogenic peptides with HLA binding motif, useful in treatment  
XX and diagnosis of cancers and viral diseases

XX  
PS Claim 1; Page 116; 150pp; English.

CC AAV45390 to AAV48214 represent specifically claimed immunogenic peptides  
CC having a human major histocompatibility complex (MHC) Class I (also  
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic  
CC peptides can bind to a specific HLA allele (i.e., HLA-A subtypes  
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell  
CC response against the antigen from which the peptide is derived.  
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are  
CC normally induced by an antigen in the form of a peptide fragment bound  
CC to a HLA molecule, rather than the intact foreign antigen itself, and  
CC are particularly important in tumour rejection and in fighting viral  
CC infections. The peptides are therefore useful therapeutically to treat  
CC or prevent viral infections and cancers in mammals (especially humans)  
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.  
CC They can be administered as vaccines to elicit an immune response in  
CC individuals susceptible or otherwise at risk of viral infection or  
CC cancer, or used to treat chronic or acute conditions. They are also  
CC useful diagnostically, and can be used to induce a cytotoxic T cell  
CC response, by contacting a cytotoxic T cell with the peptide e.g. to  
CC produce CTLs ex vivo for infusion back into a patient. The

CC polynucleotides encoding the immunogenic peptides are also useful  
CC therapeutically and for immunisation as above.

XX  
SQ Sequence 9 AA;

Query Match 91.4%; Score 53; DB 20; Length 9;  
Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYMAV 9  
111111111  
Db 1 KTWGQYMAV 9

## RESULT 12

AAV33172  
ID AAV33172 standard; peptide: 9 AA.

XX  
AC AAV33172;

XX  
DT 17-NOV-1999 (first entry)

XX  
DE Human gp100-Pmel17 peptide #1.

XX  
KW Human; protein delivery; Yersinia sp.; effector gene; mutant; antigen;  
KW immune response; cytotoxic T-lymphocyte; CTL; vaccination; treatment;  
KW pathological disorder; gp100-Pmel17.

XX  
OS Homo sapiens.

XX  
PN WO945098-A2.

XX  
PD 10-SEP-1999.

XX  
PF 03-MAR-1999; 99WO-IB00587.

XX  
PR 06-MAR-1998; 98US-0036582.

XX  
PA (VBRU/) VAN DER BRUGGEN P B.

XX  
PA (CORNU/) CORNELIS G R.

XX  
PA (BOLA/) BOLAND A M.

XX  
PA (BOON/) BOON-FALEUR T R.

XX  
PI Van Der Bruggen PB, Cornelis GR, Boland AM, Boon-Falleur TR;

XX  
DR WPI: 1999-540840/45.

XX  
PT New mutant Yersinia strains useful for treating a pathological disorder  
XX  
PS Example 1; Page 71; 80pp; English.

CC This invention describes a novel mutant Yersinia (Y1) strain, comprising  
CC mutation(s) in effector-encoding gene(s) and deficient in the production  
CC of functional effector protein(s). The invention describes (1) a  
CC quintuple mutant Yersinia strain, having the designation Yersinia  
CC enterocolitica YOPHOMP or Yersinia pseudotuberculosis YOPHAOU; (2) an  
CC expression vector (EV1) for delivering a heterologous protein into a  
CC eukaryotic cell, comprising in the 5'-3' direction: (3) a Yersinia or  
CC mutant Yersinia strain for delivering a heterologous protein into a  
CC eukaryotic cell, comprising contacting the cell with a Y1 transformed  
CC with the above vector (Y1-EV1); (4) a method for delivering a  
CC heterologous protein into a eukaryotic cell, comprising contacting the  
CC cell with a Y1 transformed with the above vector (Y1-EV1); (5) a method  
CC for inducing an immune response specific for a heterologous protein; (6)  
CC a method for inducing a cytotoxic T-lymphocyte (CTL) response specific  
CC for a heterologous protein; (7) a method for determining the efficacy of  
CC an antigen vaccination regimen in a subject. Y1 is used to treat a  
CC pathological disorder, by providing recombinant Yersinia for the safe  
CC delivery of proteins into eukaryotic cells. AAV33147-Y33178 are  
CC human-derived peptides used to illustrate the method of the invention.

XX  
SQ Sequence 9 AA;

Query Match 91.4%; Score 53; DB 20; Length 9;  
Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Caps 0;

Qy 1 KTWGQYNAV 9  
Db 1 ktwgqywav 9

## RESULT 13

AAV40211  
ID AAV40211 standard; peptide: 9 AA.

AC AAV40211;

DE 19-NOV-1999 (first entry)

XX Amino acid sequence of a human melanoma epitope.

XX Cytotoxic T cell; T lymphocyte; CD8+ epitope; T helper cell;

XX CD4+ epitope; B epitope; lipopeptide; interferon gamma; adjuvant;

XX vaccine; tumor; infection; immune response; cytokine profile;

XX autoacquired immune deficiency syndrome; papilloma; cancer; hepatitis;

XX autoimmune disease.

XX Homo sapiens.

XX FR2774687-A1.

XX 13-AUG-1999.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

XX (INSP ) INST PASTEUR LILLE.

XX Thiam K, Guillet JG, Ver Waerde C, Aurault C, Gras MH, Lohng E;

XX WPI; 1999-510734/43.

XX New lipopeptide comprising C-terminal interferon-gamma fragment with

XX attached lipophilic groups, used as interferon mimic, e.g. for treating

XX cancer or virus infection

XX Disclosure; Page 35; 53pp; French.

XX AAV40123-Y40379 represent epitopes that are able to activate cytotoxic

XX T lymphocytes (CD8+ epitopes), T helper cells (CD4+ epitopes), or

XX B epitopes recognized by corresponding antibodies. The epitopes may be

XX used in the composition of the invention. The specification describes a

XX lipopeptide that has a peptide part derived from mammalian interferon

XX gamma (IFN $\gamma$ ) and one or more lipophilic parts comprising a linear or

XX branched, (un)saturated 4-20C hydrocarbon chain or a steroid. The

XX lipopeptide mimics the activity of IFN $\gamma$ . Compositions comprising the

XX lipopeptide are used to treat or prevent any condition that responds

XX to IFN $\gamma$ , and as adjuvant for vaccines (particularly those directed

XX against tumors, viral or parasitic infections), to stimulate or

XX (re)orient the immune response between types 1 and 2 cytokine profiles.

XX Particular applications are treatment of infections (particularly

XX viral, e.g. acquired immune deficiency syndrome, papilloma (cancer) and

XX hepatitis, but also bacterial, fungal, parasitic or helminth); cancers

XX (particularly of kidney, cutaneous T cells or ovary, chronic

XX myelogenous leukemia or mesothelioma), allergy; and autoimmune

XX diseases.

XX Sequence 9 AA;

XX Query Match 91.4%; Score 53; DB 20; Length 9;

Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Caps 0;

Qy 1 KTWGQYNAV 9  
Db 1 ktwgqywav 9

## RESULT 14

AAV26867  
ID AAV26867 standard; peptide: 9 AA.

AC AAV26867;

DE 14-SEP-1999 (first entry)

XX Melanoma-derived lipopeptide epitope #8 for mixed micelles.

XX Micelle; microaggregate; induction; immune response; lipopeptide; CTL;

XX cytotoxic T-lymphocyte; epitope; lipid; helper T-lymphocyte; HTL; HBV;

XX tetanus; toxin; vaccine; HIV; hepatitis B virus; papilloma virus; p53;

XX melanoma; Plasmodium falciparum; malaria.

XX Synthetic.

XX Homo sapiens.

XX FR2771640-A1.

XX 04-JUN-1999.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX (CNRS ) CNRS CENT NAT RECH SCI.

XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

XX (INSP ) INST PASTEUR LILLE.

XX Bossus M, Bourgaud VI, Gras-Masse H, Guillet JG, Lippens G;

XX Tartar A, Wlarszeski JM;

XX WPI; 1999-349509/30.

XX Immunogenic lipopeptide micelles - comprising lipopeptides

XX containing cytotoxic and helper T-lymphocyte epitopes

XX Disclosure; Page 37; 60pp; French.

XX The invention relates to the generation of mixed micelles or

XX microaggregates for inducing an immune response comprise: (a) a first

XX lipopeptide comprising at least one CTL (cytotoxic T-lymphocyte) epitope

XX and at least one lipid unit; and (b) a second lipopeptide comprising at

XX least one HTL (helper T-lymphocyte) epitope and at least one lipid unit

XX different from that of the first lipopeptide. This peptide represents

XX an example of a lipopeptide epitope used in the invention and is derived

XX from a human melanoma protein. The immunogenic lipopeptide micelles

XX are used in vaccines, especially against HIV, hepatitis B virus (HBV),

XX papilloma viruses, p53, melanoma or Plasmodium falciparum malaria.

XX Sequence 9 AA;

XX Query Match 91.4%; Score 53; DB 20; Length 9;

XX Best Local Similarity 88.9%; Pred. No. 3.4e+05;

XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Caps 0;

XX Qy 1 KTWGQYNAV 9

XX Db 1 ktwgqywav 9

XX RESULT 15

XX AAV01753

ID AAY01753 standard; Peptide; 9 AA.  
 XX  
 AC AAY01753;  
 XX  
 XX  
 DT 25-JUN-1999 (first entry)  
 XX  
 DE Exemplary antigenic peptide derived from gp100(pme1117).  
 XX  
 DE MAGC-3; tumour associated gene; human leucocyte antigen Class II;  
 KW autologous CD4+ cell; MAGC-3 related disease; cancer; melanoma;  
 KW osteosarcoma; leukemia; carcinoma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9914326-A1.  
 XX  
 PD 25-MAR-1999.  
 XX  
 PF 04-SEP-1998; 98WO-US18601.  
 XX  
 PR 12-SEP-1997; 97US-0928615.  
 XX  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 PA (UYR-) UNIV VRIJE BRUSSEL.  
 XX  
 PI Boon-Falleur T, Chaux P, Corthals J, Heiman C;  
 PI Luiten R, Stroobant V, Thielemans K, Van Der Bruggen P;  
 DR WPI: 1999-244031/20.  
 XX  
 XX Isolated peptides that bind to human leucocyte antigen class II  
 PT molecules  
 PS  
 PT Disclosure; Page 29; 88pp; English.  
 XX  
 CC The present sequence represents an exemplary tumour associated peptide  
 CC antigen. The specification describes a MAGC-3 tumour associated gene.  
 CC Peptides (AAY01721-25) that bind human leucocyte antigen (HLA) Class II  
 CC molecules can be derived from the MAGC-3 protein. These peptides and  
 CC autologous CD4+ cells that bind to a complex of MAGC-3 peptide  
 CC and HLA Class II, are used to treat MAGC-3 related diseases,  
 CC particularly cancers (e.g. melanoma, osteosarcoma, leukemia and  
 CC various forms of carcinoma). The peptides are also used to produce  
 CC specific antibodies. Detection of of the peptides, e.g. in binding  
 CC assays, particularly with antibodies, is used for diagnosis of such  
 CC diseases.  
 CC  
 SO Sequence 9 AA;

Query Match 91.4%; Score 53; DB 20; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYWAY 9  
 DB 1 ktwgqyway 9  
 |||||

RESULT 16  
 AAY00715  
 ID AAY00715 standard; peptide; 9 AA.  
 XX  
 AC AAY00715;  
 XX  
 DT 12-MAY-1999 (first entry)  
 XX  
 DE Tumour antigen booster peptide gp100Pme1117 HLA-A2 #1.  
 XX  
 KW Tumour antigen; booster peptide; immune response modulation; allergy;  
 KW immune response enhancer; tumour cell; tumour rejection antigen;  
 KW leucocyte antigen-presenting molecule; autoimmune disease;  
 KW allograft rejection.

XX  
 OS Homo sapiens.  
 XX  
 PN WO9858956-A2.  
 XX  
 PD 30-DEC-1998.  
 XX  
 PF 19-JUN-1998; 98WO-US12894.  
 XX  
 PR 23-JUN-1997; 97US-0880979.  
 XX  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 PA Boon-Falleur T, Uytenhove C, Warner G;  
 DR WPI: 1999-105612/09.  
 XX  
 PT Immunization methods using viruses expressing antigen for priming  
 PT and booster immunizations - useful for modulating immune responses  
 PT against antigen, e.g. enhancing immune response against tumour cells  
 PT expressing tumour rejection antigens  
 PS  
 PT Disclosure; Page 10; 33pp; English.  
 XX  
 CC This sequence represents a tumour antigen booster peptide that can be  
 CC used in the method of the invention. The method is for for modulating an  
 CC immune response in a mammal against an antigen, and comprises:  
 CC (A) inducing an immune response by: (i) administering a virus containing  
 CC a nucleic acid molecule encoding the antigen or its precursor to generate  
 CC an immune response; and (ii) administering at least one booster dose  
 CC comprising a peptide including the antigen, in an adjuvant, in a combined  
 CC amount effective to enhance the initial immune response; or  
 CC (B) reducing an immune response as defined for (A) but using a  
 CC non-adjuvant with the peptide which includes the antigen, in an amount  
 CC effective to reduce the initial immune response. Method (A) is used to  
 CC enhance the immune response against tumour cells expressing tumour  
 CC rejection antigens, and against pathogens in subjects having human  
 CC leucocyte antigen-presenting molecules. Method (B) is used to reduce the  
 CC immune response in allergy, autoimmune disease, and allograft rejection.  
 CC Method (A) provides an immunisation method which, unlike prior art, is  
 CC not limited by the host immune response against viral vectors.  
 CC  
 SO Sequence 9 AA;

Query Match 91.4%; Score 53; DB 20; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYWAY 9  
 DB 1 ktwgqyway 9  
 |||||

RESULT 17  
 AAY10449  
 ID AAY10449 standard; Peptide; 9 AA.  
 XX  
 AC AAY10449;  
 XX  
 DT 12-MAY-1999 (first entry)  
 XX  
 DE HLA Class I motif peptide SEQ ID NO:379.  
 XX  
 KW Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;  
 KW immunisation; tumour; infectious disease; immunotherapy; cancer;  
 KW malignant melanoma; viral disease; hepatitis; AIDS.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9902183-A2.

PD 21-JAN-1999.  
XX  
PI 10-JUL-1998; 98WO-US14289.  
XX  
PR 10-DEC-1997; 97US-0988320.  
XX  
PR 10-JUL-1997; 97CA-2209815.  
XX  
PA (CTL-) CTL IMMUNOTHERAPIES CORP.  
XX  
PI Kuendig TM, Simard JTL;  
XX  
DR WPI; 1999-120514/10.  
XX  
PT Inducing a cytotoxic T lymphocyte response - by maintaining a level  
PT of antigen in the lymphatic system of a mammal so as to provide a  
PT sustained CTL response, used to treat, e.g. AIDS  
XX  
PS Disclosure: Page 40; 199pp; English.  
XX  
CC The present invention describes a method of inducing and/or sustaining  
CC an immunological cytotoxic T lymphocyte (CTL) response in a mammal. The  
CC method comprises: (a) delivering an antigen to the mammal at a level to  
CC induce an immunological CTL response in the mammal; and (b) maintaining  
CC the level of the antigen in the mammal's lymphatic system to maintain  
CC the immunologic CTL response. The method can be used for the delivery of  
CC e.g. a differentiation antigen, a tumour-specific multilineage antigen,  
CC an embryonic antigen, an oncogene antigen, a mutated tumour-suppressor  
CC gene antigen, or a viral antigen. They can be used for the treatment of  
CC disease such as cancer, e.g. malignant melanoma or infectious disease,  
CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery  
CC to the lymphatic system provides for potent CTL stimulation that takes  
CC place in the milieu of the lymphoid organ, and it sustains stimulation  
CC that is necessary to keep CTL active, cytotoxic and recirculating  
CC through the body. AAY10071 to AAY10639 represent examples of peptide  
CC antigens given in the present invention.  
XX  
SQ Sequence 9 AA:  
XX  
Query Match 91.4%; Score 53; DB 20; Length 9;  
Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KTWGQYMAV 9  
DB 1 ktwgqymav 9  
XX  
RESULT 18  
AAB33662  
ID AAB33662 standard; Peptide: 9 AA.  
XX  
AC AAB33662;  
XX  
DT 26-JAN-2001 (first entry)  
XX  
DE MHC class I associated immunogenic peptide SEQ ID 61.  
XX  
DE MHC class I associated immunogenic peptide SEQ ID 61.  
XX  
KM Microparticle; nucleic acid delivery; immunogenic peptide; MHC I; MHC II;  
KM major histocompatibility complex; vaginal tissue; mucosal tissue..  
XX  
OS Unidentified.  
XX  
OS Unidentified.  
XX  
PN WO200053161-A2.  
XX  
PD 14-SEP-2000.  
XX  
PF 10-MAR-2000; 2000WO-US06578.  
XX  
XX 11-MAR-1999; 99US-0266463.  
XX 27-MAY-1999; 99US-0321346.  
XX  
PA (ZYCO-) ZYCOS INC.

XX  
PI Lunsford LB, Putnam D, Hedley ML;  
XX  
DR WPI; 2000-638130/61.  
XX  
PT Microparticles useful for administering a nucleic acid into the mucosal  
PT tissue preferably vaginal tissue of an animal, comprises a polymeric  
PT matrix, a lipid and a nucleic acid molecule -  
XX  
PS Disclosure: Page 14; 96pp; English.  
XX  
CC The present invention relates to microparticles which are less than 20  
CC microns in diameter, which comprise a polymeric matrix, a lipid and a  
CC nucleic acid molecule. The microparticle is specifically not  
CC encapsulated in a liposome and does not comprise a cell. The nucleotide  
CC sequence encodes an expression product that binds to major  
CC histocompatibility complex (MHC) type I or II molecules. Peptides  
CC AAB33602-B33647 represent MHC class II associated immunogenic peptides,  
CC and AAB33648-B33710 represent MHC class I associated immunogenic  
CC peptides. The peptides are examples of the expression products of the  
CC nucleotide sequences which can be included in the microparticles of the  
CC invention. Sequences AAB33711-B33716 represent alternative expression  
CC products and nuclear localisation signals also used in the invention. The  
CC microparticles are useful for administering a nucleic acid into the  
CC mucosal tissue preferably vaginal tissue of an animal.  
XX  
SQ Sequence 9 AA:  
XX  
Query Match 91.4%; Score 53; DB 21; Length 9;  
Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KTWGQYMAV 9  
DB 1 ktwgqymav 9  
XX  
RESULT 19  
AAB23679  
ID AAB23679 standard; Peptide: 9 AA.  
XX  
AC AAB23679;  
XX  
DT 05-JAN-2001 (first entry)  
XX  
DE Cytotoxic T lymphocyte (CTL) epitope SEQ ID NO:31.  
XX  
XX ATPase; Hsp70; heat shock protein; cytotoxic T lymphocyte; CTL;  
XX immune response; infectious disease; malaria; cytotoxic T cell;  
XX cytosarctic; immunostimulant; cellular immune response inducer;  
XX protozoacide; leukaemia; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200049041-A1.  
XX  
PD 24-AUG-2000.  
XX  
PF 18-FEB-2000; 2000WO-JF00941.  
XX  
PR 19-FEB-1999; 99JP-0041535.  
XX  
XX (SDME ) SUMITOMO ELECTRIC IND CO.  
XX  
XX Shinbara N, Udono H, Yui K;  
XX  
DR WPI; 2000-543748/49.  
XX  
PT Fused protein capable of inducing cellular immune response, useful as  
PT active ingredient for drug compositions in preventing and/or treating  
PT infectious diseases such as malaria or cancer -  
XX

PS Claim 7; Page 58; 72pp; Japanese.  
XX  
CC The present invention describes a fused protein (I) prepared from a  
CC peptide containing a CTL (cytotoxic T lymphocyte) epitope recognised by  
CC cytotoxic T cells and a protein containing the ATPase domain of a heat  
CC shock protein. Also described are: (1) a drug composition containing (I)  
CC as active ingredient; (2) a DNA encoding (I); (3) an expression vector  
CC containing the DNA of (2); and (4) a transformant which can retain the  
CC expression vector of (3). (I) has cytostatic, immunostimulant and  
CC protozoacide activities, and can be used as a cellular immune response  
CC inducer. The protein is useful as an active ingredient for drug  
CC compositions in preventing and/or treating infectious diseases such as  
CC malaria or cancer e.g. to provide systemic immunity against leukaemia.  
CC The present sequence represents a specifically claimed CTL epitope  
CC for use in a fused protein of the present invention.  
XX  
SQ Sequence 9 AA;

Query Match 91.4%; Score 53; DB 21; Length 9;  
Best Local Similarity 88.9%; Pred No. 3.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYMAV 9  
| | | | |  
Db 1 ktwgywqv 9

RESULT 20

AAB08694  
ID AAB08694 standard; Peptide; 9 AA.

XX AAB08694;

XX 02-JAN-2001 (first entry)

DE Antigenic peptide from tumour rejection antigen gp100(Pmel117).

XX Epha3; HLA class II-binding peptide; human leukocyte antigen; antigen;

KW CD4+ T lymphocyte; tumour associated gene; vaccine.

XX Homo sapiens.

PN WO200050589-A1.

PD 31-AUG-2000.

PF 18-FEB-2000; 2000WO-US04326.

XX 22-FEB-1999; 99US-0121170.

PR 08-OCT-1999; 99US-0158566.

XX (LUDW-) LUDWIG INST CANCER RES.

PI Chiari R, Coulie P, Boon-Falleur T;

DR WPI: 2000-572089/53.

XX Novel tyrosine kinase receptor. Epha3 human leukocyte antigen (HLA)

PT class II binding peptide and nucleic acid encoding the receptor, useful

PT for diagnosing and treating conditions characterized by expression of

PT Epha3 gene

XX Disclosure; Page 35; 107pp; English.

PS AAB0868-B08704 represent antigenic peptides characteristic of tumours.  
CC The peptides may be combined in vaccines with a human Epha3 HLA (human  
CC leukocyte antigen) class II-binding peptide. Epha3 antigens, when  
CC presented by an antigen presenting cell having a HLA class II molecule,  
CC effectively induce activation and proliferation of CD4+ T lymphocytes.  
CC Epha3 is a tumour associated gene. Epha3 HLA binding peptides are used  
CC for selectively enriching a population of T lymphocytes. The peptides  
CC are also used for diagnosing a disorder characterized by Epha3 or Epha3

CC HLA binding peptide expression. The peptides are also used to treat a  
CC disorder characterized by Epha3 expression. The Epha3 binding peptides  
CC are useful in producing vaccines and antibody.

XX Sequence 9 AA;

Query Match 91.4%; Score 53; DB 21; Length 9;  
Best Local Similarity 88.9%; Pred No. 3.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYMAV 9  
| | | | |  
Db 1 ktwgywqv 9

RESULT 21

AAV71520  
ID AAV71520 standard; peptide; 9 AA.

XX AAV71520;

XX 12-OCT-2000 (first entry)

DE Human gp100 Pmel117 nonapeptide-1.

XX gp100 Pmel117; human; Tumour Rejection Antigen; TRA; tumour; cancer;

KW HLA; Human Leucocyte Antigen; MHC; Major Histocompatibility Complex; CTL;

KW cytolytic T-lymphocyte; immune response stimulator; prophylaxis; therapy;

XX diagnosis; TNF; tumour necrosis factor; vaccine; cytostatic.

XX Homo sapiens.

PN WO200032769-A2.

PD 08-JUN-2000.

PF 26-NOV-1999; 99WO-IB02018.

XX 27-NOV-1998; 98GB-0026143.

XX (LUDW-) LUDWIG INST CANCER RES.

PI Huang L, Van Pel A, Brasseur F, De Plaen E, Boon T;

DR WPI: 2000-412317/35.

XX Novel polypeptides expressed in tumor cells useful for treating cancers

PT have an ability to complex with a major histocompatibility complex

PT molecule and comprises a specific unbroken amino acid sequence

XX Disclosure; Page 20; 80pp; English.

PS The patent discloses MAGE-A10 and MAGE-A8 polypeptide, nonapeptide and  
CC decapeptide sequences, that function as tumour rejection antigens  
CC (TRAS). These peptides are capable of forming a complex with major  
CC histocompatibility complex (MHC) molecule type HLA-A2.1 (Human Leucocyte  
CC Antigen), that are recognised by T-lymphocytes and elicit an immune  
CC response from cytolytic T-lymphocytes (CTL). They function as an immune  
CC response stimulator. Tumour rejection antigens are useful in prophylaxis,  
CC therapy and diagnosis of tumours and are effective in controlling or  
CC preventing tumour growth. The present sequence is the human gp100 Pmel117  
CC nonapeptide-1, that corresponds to residues 154-162 of the tumour  
CC associated gene, gp100 Pmel117 encoding protein. It can be administered  
CC to induce or enhance an immune response and is presented by HLA-A2  
CC complex. This peptide can serve as a tumour rejection antigen (TRA) and  
CC in combination with adjuvants, can produce vaccines useful for treating  
CC a variety of tumours.

XX Sequence 9 AA;

Query Match 91.4%; Score 53; DB 21; Length 9;

Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYNAV 9  
| | | | | | | |  
| | | | | | | |  
Db 1 ktwgqywgqv 9

## RESULT 22

AAB02622  
ID AAB02622 standard; Peptide: 9 AA.

XX AAB02622;

XX 18-AUG-2000 (first entry)

XX Tumour associated peptide antigen from gp100 #2.

XX MAGE-A3; HLA class II; human leukocyte antigen; antibody; vaccine;

XX cancer; human; tumour; tumour associated gene product.

XX Homo sapiens.

XX WO200020581-A1.

XX 13-APR-2000.

XX 15-SEP-1999; 99WO-US21230.

XX 05-OCT-1998; 98US-0166448.

XX (LUDW-) LUDWIG INST CANCER RES.

XX (UYVR-) UNIV VIRJE BRUSSEL.

XX Chaux P, Stroobant V, Boon-Falleur T, Van Der Bruggen P;

XX Schultz ES, Van Snick J, Leche B, Thielemans K, Cortnals J;

XX Heirman C;

XX WPI: 2000-317713/27.

XX New MAGE-A3 class II binding peptides, useful to diagnose and treat

XX tumours, are fragments of MAGE-A3 which bind to and are presented to T

XX lymphocytes by human leukocyte antigen class II molecules

XX Disclosure: Page 33; 119pp; English.

XX The present invention relates to MAGE-A3 (tumour associated gene

XX product) human leukocyte antigen (HLA) class II-binding peptides (see

XX AAB02626-802395, and AAB02633-802637). These peptides are presented to T

XX cells in the context of HLA class II molecules. The peptides stimulate

XX the activity and proliferation of CD4+ T lymphocytes. The invention also

XX includes nucleotide sequences encoding MAGE-3A peptides (see AAA37928

XX and AAA37938-A37940). The peptides and nucleotide sequences can be used

XX to create antibodies against the MAGE-A3 peptides, the antibodies,

XX peptides and nucleotide sequences can be used to create a vaccine. The

XX expression of MAGE-3, particularly cancer. The methods can also be used

XX in the diagnosis of disorders associated with MAGE-3 expression. Included

XX in the invention are other human tumour antigens (see AAB02596-802637),

XX and PCR primers used in the course of the invention (see AAA37929-A37937

XX and AAA37941-A37942).

XX Sequence 9 AA:

XX

XX

XX

## RESULT 23

AAV90803  
ID AAV90803 standard; peptide: 9 AA.

XX AAV90803;

XX 25-AUG-2000 (first entry)

XX Human leukocyte antigen A2 gp100-Pmel117 peptide SEQ ID NO:32.

XX Human leukocyte antigen; HLA-B35; binding; recognition; lysis;

XX cytolytic T cell; tyrosinase; immune response; diagnosis;

XX identification; human.

XX Homo sapiens.

XX WO200021551-A1.

XX 20-APR-2000.

XX 04-OCT-1999; 99WO-US23038.

XX 09-OCT-1998; 98US-0169717.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Ooms A, De Giovanni G, Morel S, Van Den Eynde B, Boon-Falleur T;

XX WPI: 2000-317842/27.

XX Isolated peptides, sometimes derived from tyrosinase, which bind to

XX HLA-B35 leading to recognition and lysis of the resulting complexes by

XX cytolitic T cells

XX Example 3; Page 9; 20pp; English.

XX The present invention describes isolated peptides which bind to human

XX leukocyte antigen (HLA)-B35 molecules leading to recognition and lysis

XX of the resulting complexes by cytolitic T cells. The isolated peptides

XX are sometimes derived from tyrosinase. Compositions comprising the

XX peptides of the invention can be used to generate immune responses,

XX preferably in humans, but also in non-human animals to generated immune

XX components which can then be used to treat humans or diagnostically.

XX Therapeutically, the peptides are useful in generation of cytolitic T

XX cells either in vitro or in vivo which specifically lyse pathogenic

XX cells. The peptides can also be used to identify HLA-B35 positive

XX cells, or to remove HLA-B25 positive cells from mixtures containing

XX such cells. Nucleic acid molecules encoding the peptides can be used

XX inter alia as probes to identify cells which are expressing tyrosinase.

XX The present sequence represents an HLA binding peptide used in the

XX exemplification of the present invention.

XX Sequence 9 AA:

XX

OY 1 KTWGQYNAV 9  
| | | | | | | |  
| | | | | | | |  
Db 1 ktwgqywgqv 9

## RESULT 24

AAV92299  
ID AAV92299 standard; peptide: 9 AA.

XX AAV92299;

XX 10-AUG-2000 (first entry)

XX

```

DE gp100-Pmell17 antigenic peptide epitope (residues 154-162).
XX
XX gp100-Pmell17: antigen; epitope; cytotoxic T lymphocyte; CTL; complex;
KW human leukocyte antigen; HLA.
XX
XX Homo sapiens.
OS
XX WO200020445-A2.
XX
XX 13-APR-2000.
XX
XX 15-SEP-1999; 99WO-IB01664.
XX
XX 02-OCT-1998; 98US-0165863.
XX
XX 09-APR-1999; 99US-0289350.
XX
XX (CHAU/) CHAUX P.
XX (LUT/) LUTTEN R.
XX (DEMO/) DEMOTTE N.
XX (DUFE/) DUFFOUR M.
XX (LUR/) LURQUIN C.
XX (TRAV/) TRAVERSARI C.
XX (STRO/) STROEBANT V.
XX (COR/) CORNELIS G R.
XX (BOON/) BOON-FALLEUR T.
XX (VBRU/) VAN DER BRUGGEN P.
XX
XX Chaux P, Luiten R, Demotte N, Duffour M, Lurquin C, Traversari C;
PI Stroebant V, Cornelis GR, Boon-Falleur T, Van Der Bruggen P;
PI Schultz E, Wanner G;
XX
XX WPI: 2000-303739/26.
XX
XX Isolation of cytotoxic T-lymphocytes clones by successive steps of
PT stimulation and testing of lymphocytes with antigen presenting cells
PT which present antigens derived from different expression systems
XX
XX Disclosure: Page 22; 99pp; English.
XX
XX A novel method of isolation of cytotoxic T-lymphocytes (CTL) clones
CC comprises successive steps of stimulation and testing of lymphocytes
CC with antigen presenting cells (APCs) which present antigens derived
CC from different expression systems. The CTL clones isolated recognize
CC specific antigenic peptides of proteins, preferably of the MAGE family.
CC The APC is autologous and each expression systems is different from at
CC least one of the other expression systems, therefore isolating a
CC cytotoxic T cell clone specific for the protein. The method can also be
CC used to identify an antigenic peptide epitope. Isolated CTL clones
CC specific for a peptide/human leukocyte antigen (HLA) complex are claimed.
CC The CTL cells specific for the complexes, peptides or cells which present
CC the complexes on the cell surface are useful for treating pathological
CC conditions characterized by abnormal expression of the complexes.
CC
XX
XX Sequence 9 AA:
SQ

```

Query Match 91.4%; Score 53; DB 21; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 KTWGQYNAV 9
   |||||||
DB 1 ktwgqywav 9

```

RESULT 25  
 AAY84296 standard; Peptide; 9 AA.  
 AAY84296;  
 12-JUL-2000 (first entry)

```

DE Tumour associated antigen derived from gp100-Pmell17.
XX
XX tumour rejection antigen; macrophage colony stimulating gene;
KW macrophage-colony stimulating factor; antigen presenting cell;
KW human leukocyte antigen; CD8+ cytotoxic T lymphocyte.
XX
XX Homo sapiens.
OS
XX WO200013699-A1.
XX
XX 16-MAR-2000.
XX
XX 03-SEP-1999; 99WO-US20344.
XX
XX 04-SEP-1998; 98US-0099077.
XX
XX (LUDM-) LUDWIG INST CANCER RES.
XX
XX Probst-Keppler M, Van Den Eynde B, Boon-Falleur T;
PI
XX
XX WPI: 2000-256859/22.
XX
XX Isolated polypeptide used to treat subjects having a disorder
PT characterized by expression of alternative open reading frame
PT macrophage-colony stimulating factor comprises 25 amino acid residue
PT sequence -
XX
XX Disclosure: Page 21; 74pp; English.
XX
XX AAY84270-Y84303 represent peptides which are tumour associated antigens.
CC They can be administered in conjunction with the tumour rejection antigen
CC precursor of the invention to induce anti-tumour responses. The tumour
CC rejection antigen precursor of the invention is encoded by an
CC alternative open reading frame (ORF) of human macrophage colony
CC stimulating gene. Peptides derived from the alternative ORF of
CC macrophage-colony stimulating factor, when presented by an antigen
CC presenting cell having a human leukocyte antigen (HLA) class I molecule,
CC effectively induce the activation and proliferation of CD8+ cytotoxic T
CC lymphocytes. Polypeptide and nucleic acids derived from the alternate ORF
CC of macrophage-colony stimulating factor are useful for enriching
CC selectively a population of T lymphocytes with CD8+ T lymphocytes. They
CC are also useful for diagnosing a disorder characterized by expression of
CC the polypeptide, and for identifying functional variants and mimetics.
CC
XX
XX Sequence 9 AA:
SQ

```

Query Match 91.4%; Score 53; DB 21; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 KTWGQYNAV 9
   |||||||
DB 1 ktwgqywav 9

```

RESULT 26  
 AAY82979 standard; Peptide; 9 AA.  
 AAY82979;  
 19-JUN-2000 (first entry)

gp100(Pmell17) tumour associated antigen.

Tumour; tumour associated antigen; retrovirus; antisense;  
 treatment; probe; primer; HLA; cytotoxic T-lymphocyte; cancer;  
 tests; antibody; CTL; helper T-lymphocyte; MAGE; BAGE; GAGE; RAGE;  
 Gnt-V; MUM; CDK4; beta catenin; tyrosinase; Melan-A; gp100; PRAME.  
 Homo sapiens.



PN WO200006598-A1.  
 XX 10-FEB-2000.  
 XX  
 XX 15-JUL-1999; 99WO-US16236.  
 XX  
 XX 29-JUL-1998; 98US-0124398.  
 XX  
 PA (LUDM-) LUDWIG INST CANCER RES.  
 XX  
 PI Coulle P, Boon-falleur T;  
 XX  
 DR WPI; 2000-205453/18.  
 XX  
 PT Novel nucleic acids encoding melanoma associated gene products and  
 PT their fragments and variants, useful for treating endogenous retrovirus  
 PT mediated tumors, especially melanomas  
 PS Disclosure; Page 26; 77pp: English.

CC Tumor associated disorders (e.g. endogenous retrovirus mediated  
 CC tumors, especially melanomas) can be treated or ameliorated by  
 CC administering antisense nucleic acid to reduce the expression of  
 CC tumor associated genes such as HERV-AVL3-B. Progression of  
 CC a disorder characterized by the expression of the HERV-AVL3-B  
 CC endogenous retrovirus tumor rejection antigen (ERTRA) can be  
 CC diagnosed or monitored by contacting a non-testis biological  
 CC sample with an agent that binds to the complex and determining  
 CC the interaction. A disorder can also be treated by administering  
 CC an agent that enriches the presence of HLA and HERV-AVL3-B ERTA  
 CC or by administering autologous cytotoxic T-cells sufficient to  
 CC ameliorate the disorder. Fragments of the HERV-AVL3-B coding sequence  
 CC are useful as probes or amplification primers for determining the  
 CC expression of HERV-AVL3-B genes, to express tumor associated  
 CC polypeptides in vivo and in vitro and to prepare fragments of such  
 CC polypeptides to synthesize antibodies. Antigenic peptides of  
 CC HERV-AVL3-B can be useful for generating antibodies either alone or  
 CC as fusion proteins, as components of immunoassay and for determining  
 CC the binding specificity of HLA molecules and/or cytotoxic T  
 CC lymphocyte (CTL) for HERV-AVL3-B proteins. Peptides derived from  
 CC the HERV-AVL3-B coding sequence and which are presented by MHC  
 CC molecules and recognised by CTL or helper T-lymphocytes can be  
 CC combined with peptides from other tumour rejection antigens by  
 CC preparation of hybrid nucleic acids or polypeptides to produce  
 CC polypeptides. This exemplary tumour associated peptide antigen  
 CC corresponds to amino acids 154-162 of the gp100(pmel117) polypeptide.  
 CC See also AAY82953-Y82986.  
 CC  
 XX  
 SQ Sequence 9 AA:

Query Match 91.4%; Score 53; DB 21; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYNAV 9  
 | | | | | | | |  
 Db 1 ktwgqygqv 9

RESULT 27  
 AAY56614  
 ID AAY56614 standard; peptide; 9 AA.  
 XX  
 AC AAY56614;  
 XX  
 XX 21-FEB-2000 (first entry)  
 XX  
 DE gp100-pmel-117 gene MHC molecule HLA-A2 peptide SEQ ID NO:30.  
 XX  
 KW HLA-A\*0201; human leukocyte antigen: cytolytic T cell; CTL; tumour;  
 KW melan-A; peripheral blood lymphocyte; PBL; immune complex; melanoma;  
 KW MHC molecule; beta2-microglobulin; cytotoxic T lymphocyte; vaccine;  
 KW

KW immune response; cancer; tyrosinase; tumour rejection antigen;  
 KW major histocompatibility complex.  
 XX  
 XX Synthetic.  
 OS Homo sapiens.  
 OS  
 XX WO9950637-A2.  
 XX  
 XX 07-OCT-1999.  
 PD  
 XX  
 XX 25-MAR-1999; 99WO-US06615.  
 XX  
 XX 27-MAR-1998; 98US-0049850.  
 XX  
 PA (LUDM-) LUDWIG INST CANCER RES.  
 PA (UYOX-) UNIV OXFORD.  
 XX  
 PI Romero P, Dunbar R, Valmorl D, Ogg G, Cerrotini J, Cerundolo V;  
 XX  
 DR WPI; 2000-052636/04.  
 XX

PT New isolated complex of binding partners and immune complexes  
 PT containing major histocompatibility molecules and peptide, used to  
 PT isolate and detect cytotoxic T cells, particularly directed against  
 PT cancer  
 PS Example 50; Page 64; 91pp: English.

CC The present invention describes an isolated complex (A) comprising: (1)  
 CC first and second binding partners (Bp1, Bp2); and (ii) several immune  
 CC complexes (IC) containing a major histocompatibility complex (MHC)  
 CC molecule (I), a beta2-microglobulin molecule (b2MG) and a peptide (II)  
 CC that binds specifically to (I). (A) are used for analysis of cytolytic  
 CC T cells (CTL) for characterisation of an immune response to tumours or  
 CC for monitoring vaccine trials. Particularly they are used to isolate or  
 CC detect particular CTL (especially those in tumour-infiltrated lymph  
 CC nodes), including visualisation of antigen-specific CTL and determination  
 CC if the cells have been activated by in vivo exposure to antigen. Isolated  
 CC precursor cells may be expanded in vitro to produce cells with high  
 CC tumoricidal activity, for therapeutic or diagnostic use. A method from  
 CC the present invention allows: (i) preselection of T cell clones for use  
 CC in immunotherapy according to their homing molecules; and (ii) improves  
 CC the lytic activity of T cells populations by inhibition of natural  
 CC killer cell receptors. The present sequence represents a peptide  
 CC used in the exemplification of the present invention.  
 CC  
 XX  
 SQ Sequence 9 AA:

Query Match 91.4%; Score 53; DB 21; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYNAV 9  
 | | | | | | | |  
 Db 1 ktwgqygqv 9

RESULT 28  
 AAB31354  
 ID AAB31354 standard; peptide; 9 AA.  
 XX  
 AC AAB31354;  
 XX  
 XX 20-APR-2001 (first entry)  
 XX  
 DE Exemplary antigen characteristic of tumours, derived from gp100.  
 XX  
 KW MAGE-A1; HLA; human leukocyte antigen; CD4+ T lymphocyte; cancer;  
 KW MAGE-A1 HLA class II-binding protein; vaccine.  
 KW  
 XX Homo sapiens.  
 XX

PN WO200078806-A1.  
 XX 28-DEC-2000.  
 PD 14-JUN-2000; 2000MO-US16287.  
 XX 18-JUN-1999; 99US-0336091.  
 XX (LUDW-) LUDWIG INST CANCER RES.  
 PA Van Snick J, Lethe B, Chaux P, Boon-Falleur T, Van Der Bruggen P;  
 PI WPI; 2001-102698/11.  
 DR Novel MAGE-A1 human leukocyte antigen class II peptides which bind to  
 XX PT and are presented to the class II molecules, useful for inducing immune  
 PT response and treating cancers characterized by expression of MAGE-A1 -  
 XX  
 PS Disclosure; Page 33; 78pp; English.  
 CC AAB31302-59 represent exemplary antigens which are characteristic  
 CC of tumours. They can be used to enhance the immune response of vaccines  
 CC comprising peptides derived from human MAGE-A1 HLA (human leukocyte  
 CC antigen) class II-binding protein. Peptides derived from the MAGE-A1 HLA  
 CC binding protein stimulate the activity and proliferation of CD4+ T  
 CC lymphocytes. The MAGE-A1 HLA binding protein is useful as a diagnostic  
 CC agent for diagnosis of a disorder characterized by expression of MAGE-A1.  
 CC The protein is used for treating a disorder characterized by expression  
 CC of MAGE-A1 such as cancers e.g. melanoma, squamous cell carcinomas,  
 CC colorectal carcinomas, osteosarcomas, and lymphocytic leukemias.  
 CC Peptides derived from the MAGE-A1 HLA binding protein are useful in the  
 CC production of anti-tumour vaccines.  
 XX  
 SQ Sequence 9 AA;  
 Query Match 91.4%; Score 53; DB 22; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 3.4e+05;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KTWGQYWAY 9  
 DB 1 ktwgqyway 9  
 RESULT 29  
 AAR78643  
 ID AAR78643 standard; Protein; 10 AA.  
 XX AAR78643;  
 AC  
 XX 22-JAN-1996 (first entry)  
 DT  
 XX Immunogenic peptide of melanoma associated antigen gp100.  
 DE  
 XX Melanoma; antigen; vaccine; immunogen; primer; probe; detection;  
 KW identification; tumour; gp100.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP668350-A1.  
 PD 23-AUG-1995.  
 XX  
 PF 14-FEB-1995; 95EP-0200348.  
 XX  
 PR 21-DEC-1994; 94EP-0203709.  
 PR 16-FEB-1994; 94EP-0200337.  
 XX  
 PA (ALKU ) AKZO NOBEL NV.  
 PI Adema GT, Figdor CG;  
 XX

DR WPI; 1995-284790/38.  
 DR N-PSDB; AA096055.  
 XX  
 PT Melanoma associated antigen gp100 - used in vaccines and for the  
 PT detection of tumours  
 XX  
 PS Claim 5; Page 31; 40pp; English.  
 CC Immunogenic peptides derived from the melanoma associated antigen  
 CC (see AAR78639-45) may be used in the production of vaccines.  
 CC Nucleotide sequences encoding the immunogenic peptides may be used  
 CC as primers and probes in the detection of melanoma cells. Tumour  
 CC infiltrating lymphocytes capable of binding to the melanoma  
 CC associated antigen can be cultured ex vivo and returned to melanoma  
 CC particles, and when radiolabelled, they may be used to identify  
 CC tumour deposits.  
 XX  
 SQ Sequence 10 AA;  
 Query Match 91.4%; Score 53; DB 16; Length 10;  
 Best Local Similarity 88.9%; Pred. No. 0.013;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KTWGQYWAY 9  
 DB 1 ktwgqyway 9  
 RESULT 30  
 AAR84209  
 ID AAR84209 standard; Peptide; 10 AA.  
 XX AAR84209;  
 AC  
 XX 25-APR-1996 (first entry)  
 DT  
 XX gp100 melanoma antigen immunogenic peptide (G10-154).  
 DE  
 XX gp100; melanoma antigen recognised by T-cells; MART; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 XX  
 PF 21-APR-1995; 95WO-US05063.  
 XX  
 PR 05-APR-1995; 95US-0417174.  
 PR 22-APR-1994; 94US-0231565.  
 XX  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 XX  
 DR WPI; 1995-382963/49.  
 XX  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 XX  
 PS Claim 55; Page 131; 184pp; English.  
 CC The immunogenic peptide is derived from CDNA25 (AAR84854), a  
 CC melanoma antigen derivative of gp100 (see AAR84855). The  
 CC peptide and its derivatives (see AAR84200-R84211) are used in  
 CC medicaments (vaccines) for the treatment or prevention (by  
 CC immunization) of melanoma. Antibodies against melanoma-specific  
 CC antigens and its immunogenic peptides may be used in the

CC detection and isolation of the antigen from a sample, the  
 CC detection of which is indicative of a disease state  
 CC (melanoma or metastatic melanoma).  
 XX  
 SQ Sequence 10 AA;

Query Match 91.4%; Score 53; DB 16; Length 10;  
 Best Local Similarity 88.9%; Pred. NO. 0.013;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KTWGQYNAV 9  
 |||||  
 Db 1 ktwgqywgq 9

Search completed: July 3, 2001, 10:23:58  
 Job time: 1383 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: July 3, 2001, 10:28:10 ; Search time 37.77 Seconds

(without alignments)  
31.526 Million cell updates/sec

Title: US-09-214-836-9

Sequence: 1 KTWGQYMW 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Database :

SPREMBL.16:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.podent:\*  
12: sp.unclassified:\*  
13: sp.vertebrate:\*  
14: sp.virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	43	72.9	209	2	P73834	P73834 synechocyst
2	41	69.5	73	14	Q9DWM8	Q9DWM8 human papil
3	41	69.5	458	5	O18533	O18533 schistosoma
4	40	67.8	73	14	Q9DWM4	Q9DWM4 human papil
5	40	67.8	281	2	Q9KPY3	Q9KPY3 vibrio chol
6	40	67.8	305	2	Q45818	Q45818 chloroflexu
7	40	67.8	400	2	P74474	P74474 synechocyst
8	40	67.8	498	2	O87798	O87798 pseudomonas
9	40	67.8	501	2	O07068	O07068 burkholderi
10	39	66.1	221	2	Q9K889	Q9K889 streptomyce
11	39	66.1	236	2	Q9RXG5	Q9RXG5 deinococcus
12	39	66.1	276	1	Q9Y9N5	Q9Y9N5 aeropyrum p
13	39	66.1	469	1	O50200	O50200 thermococcu
14	39	66.1	518	5	O02622	O02622 crassostrea
15	38	64.4	160	10	O23952	O23952 gossypium h
16	38	64.4	167	10	Q9XGR7	Q9XGR7 fragaria an
17	38	64.4	169	10	Q9ZP34	Q9ZP34 nicotiana t
18	38	64.4	170	10	Q9SCA6	Q9SCA6 lycopersico
19	38	64.4	191	2	Q9EW89	Q9EW89 streptomyce

20	38	64.4	203	10	O9LIB3	O9LIB3 zinnia eleg
21	38	64.4	228	2	O59962	O59962 streptomyce
22	38	64.4	232	10	P93492	P93492 pinus taeda
23	38	64.4	232	10	P93493	P93493 pinus taeda
24	38	64.4	232	10	P93494	P93494 pinus taeda
25	38	64.4	232	10	P93495	P93495 pinus taeda
26	38	64.4	232	10	O38863	O38863 arabidopsis
27	38	64.4	239	10	O9ZP31	O9ZP31 lycopersico
28	38	64.4	242	10	O9LIB1	O9LIB1 zinnia eleg
29	38	64.4	245	10	O9LIB2	O9LIB2 zinnia eleg
30	38	64.4	245	10	O9FMT1	O9FMT1 cicer ariet
31	38	64.4	246	10	P93442	P93442 oryza sativ
32	38	64.4	247	10	O82625	O82625 lycopersico
33	38	64.4	247	10	O9M517	O9M517 tripysaria
34	38	64.4	248	10	O80622	O80622 arabidopsis
35	38	64.4	249	10	O9ZP36	O9ZP36 nicotiana t
36	38	64.4	249	10	O9ZP35	O9ZP35 nicotiana t
37	38	64.4	249	10	O9M515	O9M515 tripysaria
38	38	64.4	249	10	O9LDR9	O9LDR9 arabidopsis
39	38	64.4	250	10	O39625	O39625 cucumis sat
40	38	64.4	250	10	O9FVG9	O9FVG9 lycopersico
41	38	64.4	251	10	O40636	O40636 oryza sativ
42	38	64.4	252	10	O81133	O81133 prunus arme
43	38	64.4	252	10	O9SP23	O9SP23 regnellidiu
44	38	64.4	252	10	O9PY30	O9PY30 festuca pra
45	38	64.4	252	10	O9FUM2	O9FUM2 prunus avia

## ALIGNMENTS

RESULT 1  
ID P73834 PRELIMINARY; PRT; 209 AA.  
AC P73834;  
DT 01-FEB-1997 (TRMBLrel. 02, Created)  
DT 01-FEB-1997 (TRMBLrel. 02, Last sequence update)  
DT 01-JUN-2000 (TRMBLrel. 14, Last annotation update)  
DE HYPOHETICAL 24.5 KDA PROTEIN.  
GN SLR2003.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_Taxid=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; Pubmed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirosewa M., Suglura M., Sasamoto S., Kimura T.,  
RA Hosonouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shimp S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
DR EMBL: D90910; BAA17891.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 209 AA; 24511 MW; 0AA7432F01DD888F CRC64;

Query Match 72.9%; Score 43; DB 2; Length 209;  
Best Local Similarity 75.0%; Pred. No. 12;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYMW 8  
Db 75 KTWGQYMW 82  
RESULT 2  
Q9DWM8 PRELIMINARY; PRT; 73 AA.  
ID Q9DWM8;  
AC Q9DWM8;

DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE MAJOR CAPSID PROTEIN L1 (FRAGMENT).  
 OS Human papillomavirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10566;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FA50;  
 RA Antonsson A., Hossain S., Simon M., Hansson B.G.;  
 RT "Skin HPV found on foreheads."  
 RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY009882; AAG27044.1; -  
 FT NON\_TER 1 1  
 FT NON\_TER 73 73  
 SQ SEQUENCE 73 AA; 8479 MW; FID5BC6307D39062 CRC64;

Query Match 69.5%; Score 41; DB 14; Length 73;  
 Best Local Similarity 71.4%; Pred. No. 8.5;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 WGOYMOV 9  
 Db 42 WGEYMDV 48

RESULT 3  
 ID 018533 PRELIMINARY; PRT 458 AA.  
 AC 018533;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE PREPROCAPSID C PRECURSOR (EC 3.4.14.1).  
 OS Schistosoma japonicum (Blood fluke).  
 OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;  
 OC Trematoda; Digenea; Strigeidae; Schistosomatidae; Schistosomatidae;  
 OC Schistosoma.  
 OX NCBI\_TaxID=6182;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-CHINESE;  
 RX MEDLINE=97442731; PubMed=9297696;  
 RA Brindley P.J., Kalina B.H., Dalton J.P., Day S.R., Wong J.Y.,  
 RA Smythe M.L., Mcmanus D.P.;  
 RT "Proteolytic degradation of host hemoglobin by schistosomes."  
 RL Mol. Biochem. Parasitol. 89:1-9(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CHINESE;  
 RA Hota-Jamriska L., Dalton J.P., Askov J., Day S.R., Fan J.,  
 RA Brindley P.J.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U77932; AAC32040.1; -  
 DR HSBP; P00787; ITHE.  
 DR InterPro; IPR000169; -  
 DR InterPro; IPR000668; -  
 DR Pfam; PF00112; Peptidase\_C1; 1.  
 DR PRINTS; PR00705; PAPA1N.  
 DR PROSITE; PS00139; THIOL\_PROTEASE\_CYS; 1.  
 DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; 1.  
 KW Hydrolase; Signal; Thiol protease.  
 FT SIGNAL 1 26 POTENTIAL.  
 FT CHAIN 222 458 CATHPEPSIN C.  
 SQ SEQUENCE 458 AA; 52698 MW; ADA976566C4142C CRC64;

Query Match 69.5%; Score 41; DB 5; Length 458;  
 Best Local Similarity 83.3%; Pred. No. 55;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 TWGOYV 7  
 Db 425 SMGOYV 430

RESULT 4  
 ID 09DWY4 PRELIMINARY; PRT 73 AA.  
 AC 09DWY4;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE MAJOR CAPSID PROTEIN L1 (FRAGMENT).  
 OS Human papillomavirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10566;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FA55;  
 RA Antonsson A., Hossain S., Simon M., Hansson B.G.;  
 RT "Skin HPV found on foreheads."  
 RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY009886; AAG27048.1; -  
 FT NON\_TER 1 1  
 FT NON\_TER 73 73  
 SQ SEQUENCE 73 AA; 8323 MW; 8A94051AD736407 CRC64;

Query Match 67.8%; Score 40; DB 14; Length 73;  
 Best Local Similarity 57.1%; Pred. No. 12;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 WGOYMOV 9  
 Db 42 WGEYMDV 48

RESULT 5  
 ID 09KPY3 PRELIMINARY; PRT 281 AA.  
 AC 09KPY3;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE HYPOTHETICAL PROTEIN VC2229.  
 GN VC2229.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-EL TOR N16961 / SEROTYPE O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Yamathayan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 cholerae."  
 RL Nature 406:477-483(2000).  
 DR EMBL; AE004294; AAF95373.1; -  
 DR TIGR; VC2229; -  
 SQ SEQUENCE 281 AA; 31824 MW; 7CA75AD3494DFDD0 CRC64;

Query Match 67.8%; Score 40; DB 2; Length 281;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOY 7  
 11111  
 DB 99 WGOY 103

RESULT 6  
 045818  
 ID 045818 PRELIMINARY; PRT; 305 AA.

AC 045818; 01-NOV-1996 (TREMUREL. 01, Created)  
 DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMUREL. 16, Last annotation update)  
 DE HYPOTHETICAL 33.7 KDA PROTEIN (BCHGC).  
 GN BCHGC.  
 OS Chloroflexus aurantiacus.  
 OC Bacteria; Green non-sulfur bacteria; Chloroflexaceae group;  
 OC Chloroflexaceae; Chloroflexus.  
 OX NCBI\_TaxID=1108;

RN [1]  
 RP SEQUENCE OF 224-305 FROM N.A.  
 RC STRAIN-J10-FL;  
 RX MEDLINE=94192803; PubMed=7511541;  
 RA Niedemeier G., Shiozawa J., Lottspeich F., Feick R.;  
 RT "The primary structure of two chlorosome proteins from Chloroflexus  
 RT aurantiacus";  
 RT FBS Lett. 342:61-65(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-J10-FL;  
 RA Niedemeier G.;

RT Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94192803; PubMed=7511541;  
 RA Niedemeier G., Shiozawa J.A., Lottspeich F., Feick R.G.;

RT "The primary structure of two chlorosome proteins from Chloroflexus  
 RT aurantiacus";  
 RT FBS Lett. 342:61-65(1994).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20433268; PubMed=10976061;  
 RA Xiong J., Fischer W.M., Inoue K., Nakahara M., Bauer C.E.;

RT "Molecular evidence for the early evolution of photosynthesis.";  
 RT Science 289:1724-1730(2000).  
 RL EMBL: Z34000; CA83965.1; -;  
 DR EMBL: AF288602; AAG15233.1; -;  
 DR InterPro: IPR000537; -;  
 DR Pfam: PF01040; UblA; 1.

KM Hypothetical protein.  
 SQ SEQUENCE 305 AA; 33674 MW; F990F92F2D1C2B07 CRC64;

Query Match 67.8%; Score 40; DB 2; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOY 7  
 11111  
 DB 246 WGOY 250

RESULT 7  
 P74474  
 ID P74474 PRELIMINARY; PRT; 400 AA.

AC P74474; 01-FEB-1997 (TREMUREL. 02, Created)  
 DT 01-FEB-1997 (TREMUREL. 02, Last sequence update)  
 DT 01-OCT-2000 (TREMUREL. 15, Last annotation update)  
 DE D-ALANYL-D-ALANINE CARBOXYPEPTIDASE.  
 GN SUR1924.  
 OS Synechocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

OX NCBI\_TaxID=1148;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97061201; PubMed=8905231;  
 RA Kaneke T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naito K., Okumura S.,  
 RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
 RA Tabata S.;

RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 DR EMBL: D90915; BAI8575.1; -;  
 DR MEROPS: S12.001; -;  
 KM Carboxypeptidase.

SQ SEQUENCE 400 AA; 44316 MW; 7510481820E462F CRC64;

Query Match 67.8%; Score 40; DB 2; Length 400;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOY 7  
 11111  
 DB 335 WGOY 339

RESULT 8  
 087798  
 ID 087798 PRELIMINARY; PRT; 498 AA.

AC 087798; 01-NOV-1998 (TREMUREL. 08, Created)  
 DT 01-NOV-1998 (TREMUREL. 08, Last sequence update)  
 DT 01-MAR-2001 (TREMUREL. 16, Last annotation update)  
 DE TOLUENE, O-XYLENE MONOOXYGENASE OXYGENASE SUBUNIT.  
 GN TOXA.

OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=316;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-OX1;  
 RX MEDLINE=98432776; PubMed=9758777;  
 RA Bertoni G., Martino M., Galli E., Barbieri P.;  
 RT "Analysis of the gene cluster encoding toluene/o-xylene monooxygenase  
 RT from Pseudomonas stutzeri OX1.";  
 RT Appl. Environ. Microbiol. 64:3626-3632(1998).  
 RL EMBL: AJ005663; CA06654.1; -;  
 DR InterPro: IPR003430; -;

DR Pfam: PF02332; Phenol\_Hydrol; 1.  
 KM Monooxygenase.  
 SQ SEQUENCE 498 AA; 57725 MW; E49B0C06D5B843D8 CRC64;

Query Match 67.8%; Score 40; DB 2; Length 498;  
 Best Local Similarity 75.0%; Pred. No. 85;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 TWGQWQY 9  
 11111  
 DB 368 TWGQWQY 375

RESULT 9  
 007068  
 ID 007068 PRELIMINARY; PRT; 501 AA.

AC 007068; 01-JUL-1997 (TREMUREL. 04, Created)  
 DT 01-JUL-1997 (TREMUREL. 04, Last sequence update)  
 DT 01-MAR-2001 (TREMUREL. 16, Last annotation update)  
 DE TOLUENE-3-MONOOXYGENASE OXYGENASE SUBUNIT 1.

GN TBHA.  
 OS Burkholderia cepacia (Pseudomonas cepacia).  
 OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;  
 CC Burkholderia.  
 OX NCBI\_TaxID=292;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A1;  
 RA Ma Y., Herson D.S.;  
 RL Submitted (Apr-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF001355; AAB58740.1; -;  
 DR InterPro: IPR003430; -;  
 DR Pfam: PF02332; Phenol\_Hydrolase; 1.  
 KW Monooxygenase.  
 SQ SEQUENCE 531 AA; 58100 MW; 4AAC6E496C34000 CRC64;

Query Match 67.8%; Score 39; DB 2; Length 501;  
 Best Local Similarity 75.0%; Pred. No. 86;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 TWGQYQV 9  
 DB 369 TWGQYQV 376

RESULT 10  
 O9RK89 PRELIMINARY; PRT; 221 AA.  
 AC O9RK89;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE PUTATIVE SUBSTRATE BINDING PROTEIN (EXTRACELLULAR) (BINDING-PROTEIN-DEPENDENT TRANSPORT) (FRAGMENT).  
 GN SCF1.15.  
 OS Streptomyces coelicolor.  
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Seeger K.J., Harris D.;  
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kisser H.M., Denapalte D., Eichner A., Cullum J.,  
 RA Kinoshita H., Hopwood D.A.;  
 RT "A set of ordered cosmid and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL: AL117323; CAB55535.1; -;  
 FT NON\_TER 221  
 SQ SEQUENCE 221 AA; 24497 MW; D305A7977801DA1C CRC64;

Query Match 66.1%; Score 39; DB 2; Length 221;  
 Best Local Similarity 83.3%; Pred. No. 54;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 WQGYW 8  
 DB 73 WQGYW 78

RESULT 11  
 O9RXG5 PRELIMINARY; PRT; 236 AA.  
 AC O9RXG5;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CYTOCHROME C-TYPE BIOGENESIS PROTEIN, HEME EXPORTER PROTEIN C.  
 GN DR0348.  
 OS Deinococcus radiodurans.  
 CC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.  
 OX NCBI\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RL;  
 RX MEDLINE=20036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 RA Vamathevan J.J., Lam P., McDonald L., Uitterlind T., Zaleski C.,  
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium Deinococcus  
 radiodurans R1.";  
 RL Science 286:1571-1577(1999).  
 DR EMBL: AE001895; AAF09930.1; -;  
 DR TIGR: DR0348; -;  
 DR InterPro: IPR002541; -;  
 DR InterPro: IPR003557; -;  
 DR Pfam: PF01578; CytC\_asm; 1.  
 DR PRINTS: PR01386; CCMCBIOGENSIS.  
 SQ SEQUENCE 236 AA; 26591 MW; 18E9BD37E12CE711 CRC64;

Query Match 66.1%; Score 39; DB 2; Length 236;  
 Best Local Similarity 83.3%; Pred. No. 57;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 TWGQY 7  
 DB 101 TWGQY 106

RESULT 12  
 O9Y9N5 PRELIMINARY; PRT; 276 AA.  
 AC O9Y9N5;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE 276AA LONG HYPOTHETICAL LACTOSE TRANSPORT SYSTEM PERMEASE PROTEIN.  
 GN APE2253.  
 OS Aeropyrum pernix.  
 CC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococaceae;  
 OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococaceae;  
 OX NCBI\_TaxID=56636;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K1;  
 RX MEDLINE=99310339; PubMed=10382966;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Adaki A., Kotsugi H.,  
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
 RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,  
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic hyper-thermophilic  
 crenarchaeon, Aeropyrum pernix K1.";  
 RL DNA Res. 6:83-101(1999).  
 DR EMBL: AP000064; BAA81265.1; -;  
 DR InterPro: IPR000515; -;  
 DR Pfam: PF00528; BPD\_transp; 1.



SO SEQUENCE 276 AA; 31573 MW; A944B6CBF8032631 CRC64;

## Query Match

Best Local Similarity 66.1%; Score 39; DB 1; Length 276;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTMGOYMW 9  
Db 215 KTMGOYMW 223

## RESULT 13

050200 PRELIMINARY; PRT: 469 AA.

AC 050200; PRELIMINARY; PRT: 469 AA.  
DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE AMYLASE (EC 3.2.1.1) (ALPHA-AMYLASE) (1,4-ALPHA-D-GLUCAN  
GLUCANOHYDROLASE) (GLYCOGENASE).

GN AMT.  
OS Thermococcus sp. R3.

OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus.  
NCBI\_TaxID=54421;

OX NCBI\_TaxID=54421;  
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RT3;  
RA Jones R.A., Patel B., Beacham I.R.;  
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-D-GLUCOSIDIC  
LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.

DR EMBL: AF017454; AAB87860.1; -  
DR HSSP: P06278; 1YUS.

DR InterPro: IPR000461; -  
DR Pfam: PF00128; alpha-amylase; 1.

KW Hydrolase; Glycosidase.  
SQ SEQUENCE 469 AA; 52756 MW; C3BD614PBA4DB3E2 CRC64;

## Query Match

Best Local Similarity 66.1%; Score 39; DB 1; Length 469;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTMGOYMW 7  
Db 204 KSMDOYMW 210

## RESULT 14

002622 PRELIMINARY; PRT: 518 AA.

AC 002622; PRELIMINARY; PRT: 518 AA.  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN  
GLUCANOHYDROLASE) (GLYCOGENASE) (FRAGMENT).

GN AMY.  
OS Crassostrea gigas (Pacific oyster).

OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Ostreoida;  
OC Ostreoida; Ostreidae; Crassostrea.

OX NCBI\_TaxID=29159;

RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE-DIGESTIVE GLAND;  
RA Moal J., Daniel J.Y., Le Moine S., Sellios D., Van Wormhoudt A.,  
RA Samain J.F.;

RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-D-GLUCOSIDIC  
LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.

DR EMBL: Y08370; CAA69658.1; -  
DR HSSP: P00690; 1JFH.

DR InterPro: IPR000461; -  
DR Pfam: PF00128; alpha-amylase; 1.

DR PRINTS: PR00110; ALPHAAMYLASE.  
KW Signal; Hydrolase; Glycosidase.

FT NON\_TER 1  
FT SIGNAL <1 18 POTENTIAL.  
FT CHAIN 19 518 ALPHA-AMYLASE.

SQ SEQUENCE 518 AA; 57435 MW; 2F3864914E7278AC CRC64;

## Query Match

Best Local Similarity 66.1%; Score 39; DB 5; Length 518;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTMGOYMW 7  
Db 292 KTMGOYMW 298

## RESULT 15

023952 PRELIMINARY; PRT: 160 AA.

AC 023952; PRELIMINARY; PRT: 160 AA.  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE EXPANSIN (FRAGMENT).

OS Gossypium hirsutum (Upland cotton).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
OC Malvales; Malvaceae; Gossypium.

OX NCBI\_TaxID=3635;

RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=COKER312; TISSUE=COTTON FIBER;  
RX MEDLINE=97294938; PubMed=9150611;

RA Shimizu Y.; Aotsuka S., Hasegawa O., Kawada T., Sakuno T., Sakai F.,  
RA Hayashi T.;

RT "Changes in levels of mRNAs for cell wall-related enzymes in growing  
cotton fiber cells."

RL Plant Cell Physiol. 38:375-378(1997).  
DR EMBL: D88415; BAA21109.1; -

DR Mendel: 24197; Goshl:1201;24197.

DR InterPro: IPR000882; -  
DR Pfam: PF01357; Pollen-allergen; 1.

DR PRINTS: PR01225; EXPANSINFAMILY.  
FT NON\_TER 1  
SQ SEQUENCE 160 AA; 17672 MW; A7CF0DBDB3B5469D CRC64;

## Query Match

Best Local Similarity 64.4%; Score 38; DB 10; Length 160;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTMGOYMW 8  
Db 110 RNWGOYMW 117

## RESULT 16

09XGR7 PRELIMINARY; PRT: 167 AA.

AC 09XGR7; PRELIMINARY; PRT: 167 AA.  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE EXPANSIN (FRAGMENT).

GN EXPL.  
OS Fragaria ananassa (Strawberry).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

OC Rosales; Rosaceae; Fragaria.  
OX NCBI\_TaxID=3747;

[1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-CV. CHANDLER;  
 RX MEDLINE-97303339; PubMed-9159182;  
 RA Rose J.K.C., Lee H.H., Bennett A.B.;  
 RT "Expression of a divergent expansin gene is fruit-specific and  
 RL ripening-regulated."  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:5955-5960(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. CHANDLER;  
 RA Rose J.K.C., Lee H.H., Bennett A.B.;  
 RL Submitted (Jun-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. CHANDLER;  
 RA Rose J.K.C., Lee H.H., Bennett A.B.;  
 RL Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF163812; AAD44345.2; -.  
 DR InterPro; IPR000882; -.  
 DR Pfam; PF01357; Pollen\_allergen; 1.  
 DR PRINTS; PR01225; EXPANSNFAMILY.  
 DR ProDom; PD002179; -; 1.  
 FT NON\_TER 1 167  
 FT NON\_TER 167  
 SQ SEQUENCE 167 AA; 17947 MW; A75F1D3C715419E7 CRC64;

Query Match 64.4%; Score 38; DB 10; Length 167;  
 Best Local Similarity 62.5%; Pred. No. 58;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGOYMW 8  
 Db 160 RNMGQNMW 167

RESULT 17  
 O9P34 PRELIMINARY; PRT; 169 AA.  
 AC O9P34;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE ALPHA-EXPANSIN PRECURSOR (FRAGMENT).  
 GN NT-EXP6.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
 OC Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. BY2;  
 RA MEDLINE-99026292; PubMed-9808735;  
 RA Link B.M., Cosgrove D.J.;  
 RT "Acid-growth response and alpha-expansins in suspension cultures of  
 RT bright yellow 2 tobacco."  
 RL Plant Physiol. 118:907-916(1998).  
 DR EMBL; AF049355; AAC96082.1; -.  
 DR Mendel; 35959; Nicotiana1201; 35959.  
 DR InterPro; IPR000882; -.  
 DR Pfam; PF01357; Pollen\_allergen; 1.  
 DR PRINTS; PR01225; EXPANSNFAMILY.  
 FT NON\_TER 1 169  
 FT NON\_TER 169  
 SQ SEQUENCE 169 AA; 19000 MW; 312B1F89916662D6 CRC64;

Query Match 64.4%; Score 38; DB 10; Length 169;  
 Best Local Similarity 62.5%; Pred. No. 58;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGOYMW 8

Db 117 RNMGQNMW 124

RESULT 18  
 O9SCA6 PRELIMINARY; PRT; 170 AA.  
 AC O9SCA6;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE EXPANSIN 18 (FRAGMENT).  
 GN Exp18.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
 OC Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. WEST VIRGINIA 106; TISSUE=FRUIT;  
 RA Lemaire-Chamley M., Petit J., Causse M., Raymond P., Chevallier C.;  
 RT "Isolation and characterization of cDNAs expressed during early  
 RT development of tomato fruit by mRNA differential display."  
 RL Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ270960; CAB65694.1; -.  
 DR InterPro; IPR000882; -.  
 DR Pfam; PF01357; Pollen\_allergen; 1.  
 DR PRINTS; PR01225; EXPANSNFAMILY.  
 DR ProDom; PD002179; -; 1.  
 FT NON\_TER 1 170  
 FT NON\_TER 170  
 SQ SEQUENCE 170 AA; 18538 MW; 905131AD9496D17C CRC64;

Query Match 64.4%; Score 38; DB 10; Length 170;  
 Best Local Similarity 62.5%; Pred. No. 59;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGOYMW 8  
 Db 163 RNMGQNMW 170

RESULT 19  
 O9EW89 PRELIMINARY; PRT; 191 AA.  
 AC O9EW89;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE BETA-XYLANASE (EC 3.2.1.8) (FRAGMENT).  
 GN GXVNI.  
 OS Streptomyces olivaceoviridis (Streptomyces corchorus11).  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1921;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-E-86;  
 RA Bin Y.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-E-86;  
 RA Zhang H., Yao B., Wang Y.;  
 RT "Cloning and expression of the 23kD beta-xylanase gene from  
 RT Streptomyces olivaceoviridis E-86."  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ292317; CAC19491.1; -.  
 KW Xylan degradation; Hydrolase; Glycosidase.  
 FT NON\_TER 1 191  
 FT NON\_TER 191  
 SQ SEQUENCE 191 AA; 20781 MW; 7E1C45ADE1B6B9C CRC64;

Query Match 64.4%; Score 38; DB 2; Length 191;  
 Best Local Similarity 66.7%; Pred. No. 66;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTWGQYMW 9  
 11:1111

Db 131 KTFNQYMSV 139

RESULT 20  
 Q9ILB3 PRELIMINARY; PRT; 203 AA.

AC Q9ILB3; 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)  
 DE EXPANSIN (FRAGMENT).  
 GN EXPL.

OS Zinnia elegans.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae;  
 OC eunasterids II; Asterales; Asteraceae; Asteroidae; Heliantheae;  
 OC Zinnia.  
 OX NCBI\_TaxID=34245;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20317189; PubMed=10859177;  
 RT "Im K.H., Cosgrove D.J., Jones A.M.;  
 "Subcellular localization of expansin mRNA in xylem cells.";  
 RL Plant Physiol. 123:463-470(2000).  
 DR InterPro: IPR000882; -;  
 DR Pfam: PF01357; Pollen\_allergen; 1.  
 DR PRINTS: PR01225; EXPANSINFAMILY.  
 DR NON\_TER 1  
 FT 1  
 SQ SEQUENCE 203 AA; 22692 MW; 7DC08C7DE9795C16 CRC64;

Query Match 64.4%; Score 38; DB 10; Length 203;  
 Best Local Similarity 62.5%; Pred. No. 70;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTWGQYMW 8  
 11:1111

Db 151 RNWQNMW 158

RESULT 21  
 ID Q59962 PRELIMINARY; PRT; 228 AA.

AC Q59962; 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)  
 DE ENDO-1,4-BETA-XYLANASE (EC 3.2.1.8) (1,4-BETA-D-XYLAN  
 XYLANOXYDROLASE).  
 GN XYL.

OS Streptomyces sp.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxID=19311;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=538;  
 RA Georis J.C.E.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases  
 CC LINKAGES IN XYLANS.  
 CC EMBL: X98518; CAA67143.1; -;  
 DR HSSP: P09850; 1XNB.  
 DR InterPro: IPR001137; -;

DR Pfam: PF00457; Glyco\_hydro.11; 1.  
 DR PRINTS: PR00911; GLYDRLASE1.  
 DR PROSITE: PS00776; GLYCOSYL\_HYDROL\_F11\_1; 1.  
 DR PROSITE: PS00777; GLYCOSYL\_HYDROL\_F11\_2; 1.  
 KW xylan degradation; Hydrolase; Glycosidase.  
 FT CHAIN 39 228 ENDO-1,4-BETA-XYLANASE.  
 SQ SEQUENCE 228 AA; 24493 MW; EAA15233052E74AA CRC64;

Query Match 64.4%; Score 38; DB 2; Length 228;  
 Best Local Similarity 66.7%; Pred. No. 79;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTWGQYMW 9  
 11:1111

Db 169 KTFNQYMSV 177

RESULT 22  
 P93492 PRELIMINARY; PRT; 232 AA.

AC P93492; 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1997 (TREMBlrel. 03, last sequence update)  
 DE EXPANSIN (FRAGMENT).  
 OS Pinus taeda (loblolly pine).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Coniferopsida; Coniferales; Pinaceae; Pinus.  
 OX NCBI\_TaxID=3352;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HYPOCOTYL.  
 RA Hutchison K.W., Singer P.B., Diaz-Sala C., Greenwood M.S.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U64890; AAB40634.1; -;  
 DR Mendel: 12160; Pinata; 1201; 12160.  
 DR InterPro: IPR000882; -;  
 DR Pfam: PF01357; Pollen\_allergen; 1.  
 DR PRINTS: PR01225; EXPANSINFAMILY.  
 DR PRINTS: PR00829; LOLPALDERGN.  
 DR NON\_TER 1  
 FT 1  
 SQ SEQUENCE 232 AA; 24734 MW; 1340AF483DE0741F CRC64;

Query Match 64.4%; Score 38; DB 10; Length 232;  
 Best Local Similarity 62.5%; Pred. No. 81;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTWGQYMW 8  
 11:1111

Db 182 RNWQNMW 189

RESULT 23  
 P93493 PRELIMINARY; PRT; - 232 AA.

AC P93493; 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1997 (TREMBlrel. 03, last sequence update)  
 DE EXPANSIN (FRAGMENT).  
 OS Pinus taeda (loblolly pine).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Coniferopsida; Coniferales; Pinaceae; Pinus.  
 OX NCBI\_TaxID=3352;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HYPOCOTYL.  
 RA Hutchison K.W., Singer P.B., Diaz-Sala C., Greenwood M.S.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U64891; AAB40635.1; -;  
 DR Mendel: 12161; Pinata; 1201; 12161.

DR InterPro: IPR000882; -  
 DR Pfam: PF01357; Pollen\_allergen; 1.  
 DR PRINTS: PRO1225; EXPANSNFAMILY.  
 DR PRINTS: PRO0829; LOPLIALLERGN.  
 FT NON\_TER 1  
 SQ SEQUENCE 232 AA; 24745 MW; 1340B0573DE907CF CRC64;

Query Match  
 Best Local Similarity 64.4%; Score 38; DB 10; Length 232;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGQYQW 8  
 : ||| ||  
 DB 182 RNWGMQWQ 189

RESULT 24  
 P93494 PRELIMINARY; PRT; 232 AA.  
 AC P93494;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE EXPANSIN (FRAGMENT).  
 OS Pinus taeda (loblolly pine).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Coniferopsida; Coniferales; Pinaceae; Pinus.  
 OX NCBI\_TaxID=3352;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HYPOCOTYL;  
 RA Hutchison K.W., Singer P.B., Diaz-Sala C., Greenwood M.S.;  
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: U64892; AAB40636.1; -  
 DR Mendel: 12162; PintA:1201;12162.  
 DR InterPro: IPR000882; -  
 DR Pfam: PF01357; Pollen\_allergen; 1.  
 DR PRINTS: PRO1225; EXPANSNFAMILY.  
 DR PRINTS: PRO0829; LOPLIALLERGN.  
 FT NON\_TER 1  
 SQ SEQUENCE 232 AA; 24681 MW; 01F29E4E8774F1A0 CRC64;

Query Match  
 Best Local Similarity 64.4%; Score 38; DB 10; Length 232;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGQYQW 8  
 : ||| ||  
 DB 182 RNWGMQWQ 189

RESULT 25  
 P93495 PRELIMINARY; PRT; 232 AA.  
 AC P93495;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE EXPANSIN (FRAGMENT).  
 OS Pinus taeda (loblolly pine).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Coniferopsida; Coniferales; Pinaceae; Pinus.  
 OX NCBI\_TaxID=3352;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HYPOCOTYL;  
 RA Hutchison K.W., Singer P.B., Diaz-Sala C., Greenwood M.S.;  
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: U64893; AAB40637.1; -  
 DR Mendel: 12163; PintA:1201;12163.  
 DR InterPro: IPR000882; -

DR Pfam: PF01357; Pollen\_allergen; 1.  
 DR PRINTS: PRO1225; EXPANSNFAMILY.  
 DR PRINTS: PRO0829; LOPLIALLERGN.  
 FT NON\_TER 1  
 SQ SEQUENCE 232 AA; 24694 MW; 9238EACD9E1FC5E CRC64;

Query Match  
 Best Local Similarity 64.4%; Score 38; DB 10; Length 232;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGQYQW 8  
 : ||| ||  
 DB 182 RNWGMQWQ 189

RESULT 26  
 Q38863 PRELIMINARY; PRT; 237 AA.  
 AC Q38863;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE EXPANSIN ATX1 (FRAGMENT).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96016146; PubMed=7568110;  
 RA Sheeharhan T.Y., Shi J., Durachko D.M., Gullitlan M.J.;  
 RA McQueen-Mason S.J., Shieh M., Cosgrove D.J.;  
 RT "Molecular cloning and sequence analysis of expansins--a highly  
 RT conserved, multigene family of proteins that mediate cell wall  
 RT extension in plants.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:9245-9249(1995).  
 DR EMBL: U30476; AAB38070.1; -  
 DR Mendel: 6669; Arabid:1201;6669.  
 DR InterPro: IPR000882; -  
 DR Pfam: PF01357; Pollen\_allergen; 1.  
 DR PRINTS: PRO1225; EXPANSNFAMILY.  
 FT NON\_TER 1  
 SQ SEQUENCE 237 AA; 25155 MW; 0E5A2CB9C2943682 CRC64;

Query Match  
 Best Local Similarity 64.4%; Score 38; DB 10; Length 237;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGQYQW 8  
 : ||| ||  
 DB 186 RNWGMQWQ 193

RESULT 27  
 Q9ZP31 PRELIMINARY; PRT; 239 AA.  
 AC Q9ZP31;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE EXPANSIN PRECURSOR.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
 OC Solanales; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. 91347; TISSUE=FRUIT;

RX MEDLINE=99178803; PubMed=10080718;  
 RA Brummell D.A., Harpster M.H., Dunsnutt P.,  
 RT "Differential expression of expansin gene family members during growth  
 and ripening of tomato fruit."  
 RL Plant Mol. Biol. 39:161-169(1999).  
 DR EMBL: AF059489; AAD13633.1; -  
 DR Medel: 36058; Lycex:1201;36058.  
 DR InterPro: IPR000882; -  
 DR Pfam: PF01357; Pollen\_allergen; 1.  
 DR PRINTS: PR01225; EXPANSNFAMLY.  
 DR ProDom: PD002179; -; 1.  
 KW Signal.  
 FT CHAIN 1 20 POTENTIAL.  
 FT SIGNAL 21 239 \* EXPANSIN.  
 SQ SEQUENCE 239 AA; 25606 MW; 1C43BF3A1021788C CRC64;

Query Match 64.4%; Score 38; DB 10; Length 239;  
 Best Local Similarity 62.5%; Pred. No. 83;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KTWGQYWG 8  
 : ||| ||  
 Db 188 RNWGNMQ 195

RESULT 28  
 Q9LBB1 PRELIMINARY; PRT; 242 AA.  
 AC Q9LBB1;  
 DT 01-OCT-2000 (TREMUREL. 15, Created)  
 DT 01-OCT-2000 (TREMUREL. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMUREL. 16, Last annotation update)  
 DE EXPANSIN 3.  
 OS Zinnia elegans.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; eudicotyledons; core eudicots; Asteridae;  
 OC eusterids II; Asterales; Asteraceae; Heliantheae;  
 OC Zinnia.  
 OX NCBI\_TaxID=34245;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20317189; PubMed=10859177;  
 RA Im K.H., Cosgrove D.J., Jones A.M.,  
 RT "Subcellular localization of expansin mRNA in xylem cells."  
 RL Plant Physiol. 123:463-470(2000).  
 DR EMBL: AF230333; AAF35902.1; -  
 DR InterPro: IPR000882; -  
 DR Pfam: PF01357; Pollen\_allergen; 1.  
 DR PRINTS: PR01225; EXPANSNFAMLY.  
 DR ProDom: PD002179; -; 1.  
 SQ SEQUENCE 242 AA; 26157 MW; 168A01724FCF5B58 CRC64;

Query Match 64.4%; Score 38; DB 10; Length 242;  
 Best Local Similarity 62.5%; Pred. No. 84;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KTWGQYWG 8  
 : ||| ||  
 Db 191 RNWGNMQ 198

RESULT 29  
 Q9LBB2 PRELIMINARY; PRT; 245 AA.  
 AC Q9LBB2;  
 DT 01-OCT-2000 (TREMUREL. 15, Created)  
 DT 01-OCT-2000 (TREMUREL. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMUREL. 16, Last annotation update)  
 DE EXPANSIN 2.  
 OS Zinnia elegans.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae;  
 OC eusterids II; Asterales; Asteraceae; Heliantheae;  
 OC Zinnia.  
 OX NCBI\_TaxID=34245;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20317189; PubMed=10859177;  
 RA Im K.H., Cosgrove D.J., Jones A.M.,  
 RT "Subcellular localization of expansin mRNA in xylem cells."  
 RL Plant Physiol. 123:463-470(2000).  
 DR EMBL: AF230332; AAF35901.1; -  
 DR InterPro: IPR000882; -  
 DR Pfam: PF01357; Pollen\_allergen; 1.  
 DR PRINTS: PR01225; EXPANSNFAMLY.  
 DR ProDom: PD002179; -; 1.  
 SQ SEQUENCE 245 AA; 26201 MW; F66E26114C27E360 CRC64;

Query Match 64.4%; Score 38; DB 10; Length 245;  
 Best Local Similarity 62.5%; Pred. No. 85;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KTWGQYWG 8  
 : ||| ||  
 Db 195 RNWGNMQ 202

RESULT 30  
 Q9FNT1 PRELIMINARY; PRT; 245 AA.  
 ID Q9FNT1;  
 AC Q9FNT1;  
 DT 01-MAR-2001 (TREMUREL. 16, Created)  
 DT 01-MAR-2001 (TREMUREL. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMUREL. 16, Last annotation update)  
 DE EXPANSIN.  
 OS Cicer arietinum (Chickpea) (Garbanzo).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 OC Fabales; Fabaceae; Papilionoideae; Cicer.  
 OX NCBI\_TaxID=3827;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. CASTELLANA; TISSUE=ETIOLATED EPICOTYLS;  
 RA Dopico B., Sanchez M.A., Labrador E.,  
 RT "An expansin is expressed in chickpea epicotyls."  
 RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AJ291816; CAC19183.1; -  
 SQ SEQUENCE 245 AA; 26469 MW; 0EF6BFC835A38B23 CRC64;

Query Match 64.4%; Score 38; DB 10; Length 245;  
 Best Local Similarity 62.5%; Pred. No. 85;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 KTWGQYWG 8  
 : ||| ||  
 Db 193 RNWGNMQ 200

Search completed: July 3, 2001, 10:28:11  
 Job time: 331 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using **sw** model

Run on: July 3, 2001, 10:28:37 ; Search time 14.6 Seconds  
(without alignments)  
21.116 Million cell updates/sec

Title: US-09-214-836-9  
Perfect score: 59  
Sequence: 1 KTWGQYNOV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues  
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	59	100.0	661 1	PM17_HUMAN
2	55	93.2	626 1	PM17_MOUSE
3	50	84.7	762 1	PL15_CHICK
4	45	76.3	335 1	PTP1_YEAST
5	45	76.3	549 1	YFAQ_ECOLI
6	41	69.5	277 1	CYST_ECOLI
7	41	69.5	315 1	CCSA_CHLUV
8	41	69.5	328 1	CCSA_ARATH
9	41	69.5	340 1	APBE_BUCAP
10	40	67.8	515 1	COAT_TRSV
11	40	67.8	807 1	YCD5_ECOLI
12	39	66.1	208 1	FCP_ISOGA
13	39	66.1	794 1	YB52_MYCPN
14	38	64.4	100 1	POL_SIV3
15	38	64.4	585 1	PT2B_ARATH
16	38	64.4	586 1	YEJH_ECOLI
17	38	64.4	1009 1	YE88_METJA
18	38	64.4	1046 1	POL_SIVG
19	38	64.4	1047 1	POL_SIV1
20	38	64.4	1057 1	POL_SIV1
21	38	64.4	1061 1	POL_SIV1
22	37	62.7	280 1	UBIA_ECOLI
23	37	62.7	305 1	SLEB_ECOLI
24	37	62.7	394 1	LIP3_DROME
25	37	62.7	623 1	YFE9_YEAST
26	37	62.7	677 1	YD64_MYCPN
27	37	62.7	1034 1	POL_HV2CA
28	37	62.7	1035 1	POL_HV2R
29	37	62.7	1035 1	POL_HV2W
30	37	62.7	1035 1	POL_HV2B
31	37	62.7	1036 1	POL_HV2R
32	37	62.7	1049 1	POL_HV2G1
33	37	62.7	1055 1	POL_HV2ST

34	37	62.7	1073 1	POL_HV2D1	P1757 human immun
35	37	62.7	1142 1	POL_HV2B	P1806 human immun
36	37	62.7	2211 1	FA5_BOVIN	Q28107 bos taurus
37	37	62.7	2224 1	FA5_HUMAN	P12359 homo sapien
38	36	61.0	120 1	PA21_BUNMU	P00617 bungarus mu
39	36	61.0	120 1	PA23_BUNMU	P00619 bungarus mu
40	36	61.0	145 1	PA22_BUNMU	P00618 bungarus mu
41	36	61.0	147 1	PA24_BUNMU	P17934 bungarus mu
42	36	61.0	240 1	XYNC_STRLI	P26220 bungarus mu
43	36	61.0	251 1	VG37_BPM1	P08231 bacterioph
44	36	61.0	251 1	VG37_BPM2	P08232 bacterioph
45	36	61.0	262 1	PAAG_ECOLI	P77467 escherichia

## ALIGNMENTS

RESULT 1  
ID PM17\_HUMAN STANDARD; PRT; 661 AA.  
AC P40967; Q16565; Q14817; Q12763; Q14448;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE MELANOCYTE PROTEIN PMEL 17 PRECURSOR (MELANOCYTE LINEAGE-SPECIFIC  
ANTIGEN GP100) (MELANOMA-ASSOCIATED ME20 ANTIGEN) (ME20K/ME20S)  
DE (ME20-M/ME20-S) (95 KDA MELANOCYTE-SPECIFIC SECRETED GLYCOPROTEIN).  
GN SLIV OR PHE17 OR D12553E.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92021023; PubMed=1924386;  
RA Kwon B.S., Chittamaneni C., Kozak C.A., Copeland N.G.,  
RA Gilbert D.J., Jenkins N.A., Barton D., Francke U., Kobayashi Y.,  
RA Kim K.-K.;  
RT "A melanocyte-specific gene, Pmel 17, maps near the silver coat color  
locus on mouse chromosome 10 and is in a syntenic region on human  
chromosome 12.";  
RT Proc. Natl. Acad. Sci. U.S.A. 88:9228-9232(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94327568; PubMed=7519602;  
RA Adema G.J., de Boer A.J., Vogel A.M., Loenen W.A., Figdor C.G.;  
RT "Molecular characterization of the melanocyte lineage-specific  
antigen gp100.";  
RT J. Biol. Chem. 269:20126-20133(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96154052; PubMed=8592076;  
RA Ballin T., Lee S.T., Spritz R.A.;  
RT "Genomic organization and sequence of D1253E (Pmel 17), the human  
homologue of the mouse silver (sl) locus.";  
RT J. Invest. Dermatol. 106:24-27(1996).  
RN [4]  
RP SEQUENCE FROM N.A. AND SEQUENCE OF 25-53.  
RX MEDLINE=94235165; PubMed=8179825;  
RA Mershes G.A., Marken J.S., Neubauer M., Aruffo A., Hellstrom I.,  
RA Hellstrom K.E., Marguardt H.;  
RT "Cloning and expression of the gene for the melanoma-associated ME20  
antigen.";  
RT DNA Cell Biol. 13:87-95(1994).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX Kwon B.S., Kim K., Heng H.H., Shi X.M., Tsui L., Lee Z.H.,  
RA Yoon B., Pickard R.T.;  
RT Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RX Vogel A.;  
RT Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.

Query Match	Score	DB 1	Length	661
Best Local Similarity	100.0%			
1- FUNCTION: COULD BE A MELANOGENIC ENZYME. COULD REPRESENT AN ONCOFETAL SELF-ANTIGEN THAT IS NORMALLY EXPRESSED AT LOW LEVELS IN OUTGROWTH ADULT MELANOCYTES BUT OVEREXPRESSED BY PROLIFERATING NEONATAL MELANOCYTES AND DURING TUMOR GROWTH. RELEASE OF THE SOLUBLE FORM, ME20-S, COULD PROTECT TUMOR CELLS FROM ANTIBODY MEDIATED IMMUNITY.	100.0%			
1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL). THERE IS ALSO A SECRETED SOLUBLE FORM, ME20-S, PROBABLY PRODUCT OF PROTEOLYTIC CLEAVAGE.	100.0%			
1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOMAS. SOME EXPRESSION WAS FOUND IN DISPLASTIC NEVI. NOT FOUND IN NORMAL TISSUES NOR IN CARCINOMAS.	100.0%			
1- SIMILARITY: BELONGS TO THE PMEL-17/MMB FAMILY.	100.0%			
1- SIMILARITY: CONTAINS 1 PKD DOMAIN.	100.0%			
THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -- THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	100.0%			
EMBL; M77348; AAA60121.1; -	100.0%			
EMBL; S73003; AAC60634.1; -	100.0%			
EMBL; U31799; AAB00386.1; -	100.0%			
EMBL; U31808; AAB00386.1; JOINED.	100.0%			
EMBL; U31807; AAB00386.1; JOINED.	100.0%			
EMBL; U31797; AAB00386.1; JOINED.	100.0%			
EMBL; U31798; AAB00386.1; JOINED.	100.0%			
EMBL; U01874; AAA18479.1; -	100.0%			
EMBL; U20093; AAB19181.1; -	100.0%			
EMBL; U19491; AAB19181.1; JOINED.	100.0%			
EMBL; M32295; AAA35930.1; ALT_INT.	100.0%			
MIM; 155550; -	100.0%			
InterPro: IPR000601; -	100.0%			
Pfam: PF00801; PKD. 1.	100.0%			
PROSITE; PS50093; PKD. 1.	100.0%			
Transmembrane; Glycoprotein; Signal; Melanin biosynthesis; Repeat; Antigen.	100.0%			
CHAIN SIGNAL 1 24	100.0%			
DOMAIN 25 595	100.0%			
TRANSMEM 596 616	100.0%			
DOMAIN 617 661	100.0%			
DOMAIN 255 292	100.0%			
DOMAIN 315 444	100.0%			
REPEAT 315 327	100.0%			
REPEAT 328 340	100.0%			
REPEAT 341 353	100.0%			
REPEAT 354 366	100.0%			
REPEAT 367 379	100.0%			
REPEAT 380 392	100.0%			
REPEAT 393 405	100.0%			
REPEAT 406 418	100.0%			
REPEAT 419 431	100.0%			
REPEAT 432 444	100.0%			
REPEAT 81 106	100.0%			
CARBOHYD 106 106	100.0%			
CARBOHYD 111 111	100.0%			
CARBOHYD 321 321	100.0%			
CARBOHYD 321 321	100.0%			
CARBOHYD 568 568	100.0%			
CONFLICT 274 274	100.0%			
CONFLICT 587 587	100.0%			
CONFLICT 592 592	100.0%			
CONFLICT 597 597	100.0%			
CONFLICT 642 661	100.0%			
SEQUENCE 661 AA; 70255 MM; 8A904FAB16715653 CMC64;	100.0%			
MELANOCYTE PROTEIN PMEL. 17.	100.0%			
EXTRACELLULAR (POTENTIAL).	100.0%			
POTENTIAL.	100.0%			
CYTOPLASMIC (POTENTIAL).	100.0%			
PKD.	100.0%			
10 X 13 AA APPROXIMATE TANDEM REPEATS.	100.0%			
1.	100.0%			
2.	100.0%			
3.	100.0%			
4.	100.0%			
5.	100.0%			
6.	100.0%			
7.	100.0%			
8.	100.0%			
9.	100.0%			
10.	100.0%			
N-LINKED (GLCNAC. . .) (POTENTIAL).	100.0%			
N-LINKED (GLCNAC. . .) (POTENTIAL).	100.0%			
N-LINKED (GLCNAC. . .) (POTENTIAL).	100.0%			
N-LINKED (GLCNAC. . .) (POTENTIAL).	100.0%			
N-LINKED (GLCNAC. . .) (POTENTIAL).	100.0%			
L -> P (IN REF. 1 AND 5).	100.0%			
P -> PVGILLT (IN REF. 1).	100.0%			
G -> GG (IN REF. 1).	100.0%			
P -> R (IN REF. 1).	100.0%			
RIFSCIPGENSPILSGOV -> ASSALVPLVRIAPSSVG	100.0%			
SRSYSHMML (IN REF. 6).	100.0%			

Matches	9;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	KTWGQYMOY	9						
Db	154	KTWGQYMOY	162						
RESULT 2									
PM17_MOUSE	2								
ID	PM17_MOUSE	STANDARD;	PRT;	626	AA.				
AC	Q60696;								
DT	01-NOV-1997	(Rel. 35, Created).							
DT	01-NOV-1997	(Rel. 35, Last sequence update).							
DT	01-OCT-2000	(Rel. 40, Last annotation update)							
DE	MELANOCYTE PROTEIN PMEL17 PRECURSOR (SILVER LOCUS PROTEIN).								
GN	SILV OR PMEL17 OR D10H12553E OR S1.								
OS	Mus musculus (Mouse).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
OX	NCBI_TaxID=10090;								
RA	[1]								
RA	SEQUENCE FROM N.A.								
RP	STRAIN=C57BL/6; TISSUE=Skin;								
RX	STRIDE=95175358; PubMed=7870580;								
RA	Kwon B.S., Halaban R., Ponnazhagan S., Kim K., Chintamaneni C.,								
RA	Bennett D., Pickard R.T.;								
RT	"Mouse silver mutation is caused by a single base insertion in the								
RT	putative cytoplasmic domain of Pmel 17."								
RL	Nucleic Acids Res. 23:154-158(1995).								
CC	-1- FUNCTION: COULD BE A MELANOGENIC ENZYME.								
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).								
CC	-1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCYTES.								
CC	-1- DISEASE: DEFECTS IN PMEL17 ARE THE CAUSE OF THE SILVER COAT COLOR								
CC	WHICH SEEMS TO BE DUE TO PREMATURE DEATH OF PIGMENT CELLS DURING								
CC	THE HAIR CYCLE.								
CC	-1- SIMILARITY: BELONGS TO THE PMEL-17/MBB FAMILY.								
CC	-1- SIMILARITY: CONTAINS 1 PRO DOMAIN.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on les-								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>								
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).								
CC	-----								
DR	EMBL; U14133; AAA69538.1; -								
DR	MGD; MG1:98301; si-								
DR	PROSITE; PS50093; PKD; 1.								
DR	Transmembrane; Glycoprotein; Signal; Melanin biosynthesis; Repeat;								
KW	Disease mutation.								
FT	CHAIN	1	24						
FT	CHAIN	25	626						
FT	DOMAIN	25	562						
FT	TRANSMDM	563	583						
FT	DOMAIN	584	626						
FT	DOMAIN	255	292						
FT	DOMAIN	315	411						
FT	REPEAT	315	327						
FT	REPEAT	328	340						
FT	REPEAT	341	353						
FT	REPEAT	354	366						
FT	REPEAT	367	379						
FT	REPEAT	380	392						
FT	REPEAT	393	411						
FT	CARBOHYD	81	81						
FT	CARBOHYD	106	106						
FT	CARBOHYD	111	111						
FT	CARBOHYD	535	535						
FT	VARIANT	170	170						



FT VARIANT 603 626 AAPASGLRAGLGENSEPILSGOV -> SSASIRSSRPMP  
 FT SEQUENCE 626 AA: 65980 MW: 7AB94ID23FBI044 CRC64;  
 SQ

Query Match  
 Best Local Similarity 93.2%; Score 55; DB 1; Length 626;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
 1111111111  
 DB 154 KTWGKRYMOV 162

RESULT 3  
 P115\_CHICK STANDARD; PRT; 762 AA.  
 AC Q98917;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE MELANOSOMAL MATRIX PROTEIN 115 KDA PROTEIN PRECURSOR.  
 GN MPM115.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WHITE LECHORN; TISSUE=retinal pigment epithelium;  
 RX MEDLINE=92020667; PubMed=1924173;  
 RA Mochizuki M., Agata K., Eguchi G.;  
 RT "Complete sequence and expression of a cDNA encoding a chicken  
 RT 115-kDa melanosomal matrix protein.";  
 RT Pigment Cell Res. 4:41-47(1991).  
 RL [2]  
 RN CHARACTERIZATION.  
 RP STRAIN=WHITE LECHORN; TISSUE=retinal pigment epithelium;  
 RX MEDLINE=88311098; PubMed=3409326;  
 RA Mochizuki M., Agata K., Kobayashi H., Yamamoto T.S., Eguchi G.;  
 RT "Expression of gene coding for a melanosomal matrix protein  
 RT transcriptionally regulated in the transdifferentiation of chick  
 RT embryo pigmented epithelial cells.";  
 RL Cell Differ. 24:67-74(1988).  
 CC -1- FUNCTION: MIGHT BE REQUIRED FOR POLYMERIZATION OF MELANIN ONTO THE  
 CC CORE STRUCTURE OF MELANOSOMES WITH ENZYMIC FUNCTION OF TYROSINASE.  
 CC -1- SUBCELLULAR LOCATION: ON THE FIBROUS MATRIX STRUCTURE OF THE  
 CC PREMELANOSOME.  
 CC -1- TISSUE SPECIFICITY: SPECIFIC TO PIGMENTED EPITHELIAL CELLS AND  
 CC MELANOCYTES. NOT EXPRESSED IN LENS, NEURAL RETINA, BRAIN, HEART,  
 CC GIZZARD OR LIVER.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING THE REDIFFERENTIATION OF  
 CC PIGMENTED EPITHELIAL CELLS (PEC).  
 CC -1- PTM: GLYCOSYLATED.  
 CC -1- SIMILARITY: BELONGS TO THE PHEM-17/NMB FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 PKD DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D88348; BAA13589.1; -  
 CC InterPro: IPR000601; -  
 CC Pfam: PF00801; PKD.1;  
 CC PROSITE: PS50093; PKD.1;  
 CC Signal: Glycoprotein; Repeat.  
 CC SIGNAL 1 19 POTENTIAL.  
 CC CHAIN 20 762 MELANOSOMAL MATRIX PROTEIN 115 KDA

FT DOMAIN 223 323 PROTEIN.  
 FT REPEAT 441 532 PKD.  
 FT REPEAT 441 464 4 X 20-24 AA APPROXIMATE TANDEM REPEATS.  
 FT REPEAT 465 488 1.  
 FT REPEAT 489 508 2.  
 FT REPEAT 509 532 3.  
 FT REPEAT 511 532 4.  
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 346 346 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 659 659 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 762 AA: 77356 MW: 172C8DB4FDC7C6 CRC64;

Query Match  
 Best Local Similarity 84.7%; Score 50; DB 1; Length 762;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 TWGQYMOV 9  
 1111111111  
 DB 161 TWGKRYMOV 168

RESULT 4  
 PP1\_YEAST STANDARD; PRT; 335 AA.  
 ID PP1\_YEAST  
 AC P25044;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE PROTEIN-TYROSINE PHOSPHATASE 1 (EC 3.1.3.48) (PTPASE 1).  
 GN PTP1 OR YDL230W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91302312; PubMed=1649172;  
 RA Guan K., Deschenes R.J., Qiu H., Dixon J.E.;  
 RT "Cloning and expression of a yeast protein tyrosine phosphatase.";  
 RT J. Biol. Chem. 266:12964-12970(1991).  
 RL [2]  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Rasmussen S.W.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: IS NOT REQUIRED FOR VEGETATIVE GROWTH.  
 CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
 CC PROTEIN TYROSINE + ORTHOPHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
 CC TYROSINE PHOSPHATASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M64062; AAA34923.1; -  
 CC EMBL: Z74278; CAA98809.1; -  
 CC PIR: A39862; A39862.  
 CC HSP: P18052; 1YPO.  
 CC SGD: S0002389; PP1.  
 CC InterPro: IPR000242; -  
 CC InterPro: IPR000387; -  
 CC Pfam: PF00102; Y-Phosphatase; 1.  
 CC PRINTS: PR00700; PTPPHPTASE.  
 CC PROSITE: PS00383; TYR\_PHOSPHATASE\_1; 1.  
 CC PROSITE: PS50056; TYR\_PHOSPHATASE\_2; 1.



Query Match 69.5%; Score 41; DB 1; Length 277;  
 Best Local Similarity 62.5%; Pred. No. 9.1;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 TWGQYMW 9  
 : |||||  
 Db 43 SNAQYMEV 50

RESULT 7  
 CCSA\_CHLVU STANDARD; PRT; 315 AA.  
 AC P56315;  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE CYTOCHROME C BIOGENESIS PROTEIN CCSA.  
 GN Chlorocella vulgaris.  
 OS Chloroplast.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;  
 OC Chlorellaceae; Chlorella.  
 OX NCBI\_TaxID=3077;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IAM C-27 / TAMITA;  
 RX MEDLINE=97303241; PubMed=9159184;  
 RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,  
 RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,  
 RA Imanura A., Yoshinaga K., Sugitara M.;  
 RT "Complete nucleotide sequence of the chloroplast genome from the  
 RT green alga Chlorocella vulgaris: the existence of genes possibly  
 RT involved in chloroplast division."  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).  
 CC -1- FUNCTION: REQUIRED DURING CYTOCHROME BIOGENESIS AT THE STEP OF  
 CC HEME ATTACHMENT (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE CCMF/CYCK/CCL1/NRFE/CCSA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; AB001684; BAA57962.1; -  
 CC DR InterPro; IPR002541; -  
 CC DR Pfam; PF01578; Cytc.asm; 1.  
 CC Cytochrome c-type biogenesis; Chloroplast.  
 CC KW CYTOCHROME C-TYPE BIOGENESIS; Chloroplast.  
 CC SQ SEQUENCE 315 AA; 35471 MW; 5020388E94FE4F10 CRC64;

Query Match 69.5%; Score 41; DB 1; Length 315;  
 Best Local Similarity 71.4%; Pred. No. 10;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYMW 7  
 : |||||  
 Db 243 ETWGNW 249

RESULT 8  
 CCSA\_ARATH STANDARD; PRT; 328 AA.  
 AC P56770;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CYTOCHROME C BIOGENESIS PROTEIN CCSA.  
 GN CCSA.  
 OS Arabidopsis thaliana (Mouse-ear cress).

OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20039611; PubMed=10574454;  
 RA Sato S., Nakamura Y., Kaneko T., Asamizu E., Tabata S.,  
 RA "Complete structure of the chloroplast genome of Arabidopsis  
 RA thaliana."  
 RT DNA Res. 6:283-290(1999).  
 CC -1- FUNCTION: REQUIRED DURING CYTOCHROME BIOGENESIS AT THE STEP OF  
 CC HEME ATTACHMENT (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE CCMF/CYCK/CCL1/NRFE/CCSA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; AP000423; BAA84436.1; -  
 CC DR InterPro; IPR002541; -  
 CC DR Pfam; PF01578; Cytc.asm; 1.  
 CC Cytochrome c-type biogenesis; Chloroplast.  
 CC KW CYTOCHROME C-TYPE BIOGENESIS; Chloroplast.  
 CC SQ SEQUENCE 328 AA; 37732 MW; C8BD1508B2924D6F CRC64;

Query Match 69.5%; Score 41; DB 1; Length 328;  
 Best Local Similarity 71.4%; Pred. No. 11;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYMW 7  
 : |||||  
 Db 255 ETWGSYMW 261

RESULT 9  
 APBE\_BUCAP STANDARD; PRT; 340 AA.  
 AC O85292;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE PROBABLE THIAMINE BIOSYNTHESIS PROTEIN APBE PRECURSOR.  
 GN APBE.  
 OS Buchnera aphidicola (subsp. Schizaphis graminum).  
 CC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 CC NCB1\_TaxID=98794;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98353428; PubMed=9688822;  
 RA Thao M.L., Baumann P.;  
 RA "Sequence analysis of a DNA fragment from Buchnera aphidicola (Aphid  
 RA endosymbiont) containing the genes dapt-htra-llvi-llvh-ftsl-ftsl-  
 RA mure."  
 RT Curr. Microbiol. 37:214-216(1998).  
 [2]  
 RP REVISIONS.  
 RA Thao M.L., Baumann P.;  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: INVOLVED IN THE CONVERSION OF AMINOIMIDAZOLE RIBOTIDE  
 CC (AIR), A PURINE INTERMEDIATE, TO THE 4-AMINO-5-HYDROXYMETHYL-2-  
 CC METHYL PYRIMIDINE (HMP) MOIETY OF THIAMINE (BY SIMILARITY).  
 CC -1- PATHWAY: BIOSYNTHESIS OF THE PYRIMIDINE MOIETY OF THIAMINE.  
 CC -1- SIMILARITY: BELONGS TO THE APBE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF060492; AAC32332.2; -

KW Thiamine biosynthesis; Signal.

FT SIGNAL 1 340 POTENTIAL.

FT CHAIN ? 340 PROBABLE THIAMINE BIOSYNTHESIS PROTEIN

FT ABBE.

SEQUENCE 340 AA; 38703 MW; 70EBABCCF8754C5C CRC64;

Query Match 69.5%; Score 41; DB 1; Length 340;

Best Local Similarity 77.8%; Pred. No. 11;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTWGYWQV 9

DB 34 KTWGYWQV 42

RESULT 10

COAT\_TRSV STANDARD; PRT; 515 AA.

AC 088894;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE COAT PROTEIN PRECURSOR (CAPSID PROTEIN).

OS Tobacco ringspot virus (TobrsV) (TRSV).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;

OC Nepovirus.

OX NCBL\_TaxID=12282;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE-94152161; PubMed-8109164;

RA Buckley B., Silva S., Singh S.;

RT "Nucleotide sequence and in vitro expression of the capsid protein

gene of tobacco ringspot virus.";

RL Virus Res. 30:335-349(1993).

RN [2]

RP ERATUM.

RX MEDLINE-95274284; PubMed-7754671;

RA Buckley B., Silva S., Singh S.;

RT Virus Res. 35:111-111(1995).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).

RX MEDLINE-98179933; PubMed-9519407;

RA Chandrasekar V., Johnson J.E.;

RT "The structure of tobacco ringspot virus: a link in the evolution of

icosahedral capsids in the picornavirus superfamily.";

RL Structure 6:157-171(1998).

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE-94152161; PubMed-8109164;

RA Buckley B., Silva S., Singh S.;

RT "Nucleotide sequence and in vitro expression of the capsid protein

gene of tobacco ringspot virus.";

RL Virus Res. 30:335-349(1993).

RN [2]

RP ERATUM.

RX MEDLINE-95274284; PubMed-7754671;

RA Buckley B., Silva S., Singh S.;

RT Virus Res. 35:111-111(1995).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).

RX MEDLINE-98179933; PubMed-9519407;

RA Chandrasekar V., Johnson J.E.;

RT "The structure of tobacco ringspot virus: a link in the evolution of

icosahedral capsids in the picornavirus superfamily.";

RL Structure 6:157-171(1998).

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE-94152161; PubMed-8109164;

RA Buckley B., Silva S., Singh S.;

RT "Nucleotide sequence and in vitro expression of the capsid protein

gene of tobacco ringspot virus.";

RL Virus Res. 30:335-349(1993).

RN [2]

RP ERATUM.

RX MEDLINE-95274284; PubMed-7754671;

RA Buckley B., Silva S., Singh S.;

RT Virus Res. 35:111-111(1995).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).

RX MEDLINE-98179933; PubMed-9519407;

RA Chandrasekar V., Johnson J.E.;

RT "The structure of tobacco ringspot virus: a link in the evolution of

icosahedral capsids in the picornavirus superfamily.";

RL Structure 6:157-171(1998).

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE-94152161; PubMed-8109164;

RA Buckley B., Silva S., Singh S.;

RT "Nucleotide sequence and in vitro expression of the capsid protein

gene of tobacco ringspot virus.";

RL Virus Res. 30:335-349(1993).

RN [2]

RP ERATUM.

RX MEDLINE-95274284; PubMed-7754671;

RA Buckley B., Silva S., Singh S.;

RT Virus Res. 35:111-111(1995).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).

RX MEDLINE-98179933; PubMed-9519407;

RA Chandrasekar V., Johnson J.E.;

RT "The structure of tobacco ringspot virus: a link in the evolution of

icosahedral capsids in the picornavirus superfamily.";

RL Structure 6:157-171(1998).

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE-94152161; PubMed-8109164;

RA Buckley B., Silva S., Singh S.;

RT "Nucleotide sequence and in vitro expression of the capsid protein

gene of tobacco ringspot virus.";

RL Virus Res. 30:335-349(1993).

RN [2]

RP ERATUM.

RX MEDLINE-95274284; PubMed-7754671;

RA Buckley B., Silva S., Singh S.;

RT Virus Res. 35:111-111(1995).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).

RX MEDLINE-98179933; PubMed-9519407;

RA Chandrasekar V., Johnson J.E.;

RT "The structure of tobacco ringspot virus: a link in the evolution of

icosahedral capsids in the picornavirus superfamily.";

RL Structure 6:157-171(1998).

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE-94152161; PubMed-8109164;

RA Buckley B., Silva S., Singh S.;

RT "Nucleotide sequence and in vitro expression of the capsid protein

gene of tobacco ringspot virus.";

RL Virus Res. 30:335-349(1993).

RN [2]

RP ERATUM.

RX MEDLINE-95274284; PubMed-7754671;

RA Buckley B., Silva S., Singh S.;

RT Virus Res. 35:111-111(1995).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).

RX MEDLINE-98179933; PubMed-9519407;

RA Chandrasekar V., Johnson J.E.;

RT "The structure of tobacco ringspot virus: a link in the evolution of

icosahedral capsids in the picornavirus superfamily.";

RL Structure 6:157-171(1998).

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE-94152161; PubMed-8109164;

RA Buckley B., Silva S., Singh S.;

RT "Nucleotide sequence and in vitro expression of the capsid protein

gene of tobacco ringspot virus.";

RL Virus Res. 30:335-349(1993).

RN [2]

RP ERATUM.

RX MEDLINE-95274284; PubMed-7754671;

RA Buckley B., Silva S., Singh S.;

RT Virus Res. 35:111-111(1995).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).

RX MEDLINE-98179933; PubMed-9519407;

RA Chandrasekar V., Johnson J.E.;

RT "The structure of tobacco ringspot virus: a link in the evolution of

icosahedral capsids in the picornavirus superfamily.";

RL Structure 6:157-171(1998).

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE-94152161; PubMed-8109164;

RA Buckley B., Silva S., Singh S.;

RT "Nucleotide sequence and in vitro expression of the capsid protein

gene of tobacco ringspot virus.";

RL Virus Res. 30:335-349(1993).

RN [2]

RP ERATUM.

RX MEDLINE-95274284; PubMed-7754671;

RA Buckley B., Silva S., Singh S.;

RT Virus Res. 35:111-111(1995).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).

RX MEDLINE-98179933; PubMed-9519407;

RA Chandrasekar V., Johnson J.E.;

RT "The structure of tobacco ringspot virus: a link in the evolution of

icosahedral capsids in the picornavirus superfamily.";

RL Structure 6:157-171(1998).

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE-94152161; PubMed-8109164;

RA Buckley B., Silva S., Singh S.;

RT "Nucleotide sequence and in vitro expression of the capsid protein

gene of tobacco ringspot virus.";

RL Virus Res. 30:335-349(1993).

RN [2]

RP ERATUM.

RX MEDLINE-95274284; PubMed-7754671;

RA Buckley B., Silva S., Singh S.;

RT Virus Res. 35:111-111(1995).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).

RX MEDLINE-98179933; PubMed-9519407;

RA Chandrasekar V., Johnson J.E.;

RT "The structure of tobacco ringspot virus: a link in the evolution of

icosahedral capsids in the picornavirus superfamily.";

RL Structure 6:157-171(1998).

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE-94152161; PubMed-8109164;

RA Buckley B., Silva S., Singh S.;

RT "Nucleotide sequence and in vitro expression of the capsid protein

gene of tobacco ringspot virus.";

RL Virus Res. 30:335-349(1993).

RN [2]

RP ERATUM.

RX MEDLINE-95274284; PubMed-7754671;

RA Buckley B., Silva S., Singh S.;

RT Virus Res. 35:111-111(1995).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).

RX MEDLINE-98179933; PubMed-9519407;

RA Chandrasekar V., Johnson J.E.;

RT "The structure of tobacco ringspot virus: a link in the evolution of

icosahedral capsids in the picornavirus superfamily.";

RL Structure 6:157-171(1998).

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE-94152161; PubMed-8109164;

RA Buckley B., Silva S., Singh S.;

RT "Nucleotide sequence and in vitro expression of the capsid protein

gene of tobacco ringspot virus.";

RL Virus Res. 30:335-349(1993).

RN [2]

RP ERATUM.

RX MEDLINE-95274284; PubMed-7754671;

RA Buckley B., Silva S., Singh S.;

RT Virus Res. 35:111-111(1995).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).

RX MEDLINE-98179933; PubMed-9519407;

RA Chandrasekar V., Johnson J.E.;

RT "The structure of tobacco ringspot virus: a link in the evolution of

icosahedral capsids in the picornavirus superfamily.";

RL Structure 6:157-171(1998).

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE-94152161; PubMed-8109164;

RA Buckley B., Silva S., Singh S.;

RT "Nucleotide sequence and in vitro expression of the capsid protein

gene of tobacco ringspot virus.";

RL Virus Res. 30:335-349(1993).

RN [2]

RP ERATUM.

RX MEDLINE-95274284; PubMed-7754671;

RA Buckley B., Silva S., Singh S.;

RT Virus Res. 35:111-111(1995).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).

RX MEDLINE-98179933; PubMed-9519407;

RA Chandrasekar V., Johnson J.E.;

RT "The structure of tobacco ringspot virus: a link in the evolution of

icosahedral capsids in the picornavirus superfamily.";

RL Structure 6:157-171(1998).

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE-94152161; PubMed-8109164;

RA Buckley B., Silva S., Singh S.;

RT "Nucleotide sequence and in vitro expression of the capsid protein

gene of tobacco ringspot virus.";

RL Virus Res. 30:335-349(1993).

RN [2]

RP ERATUM.

Db 314 WGYW 318

RESULT 12  
FCP\_ISOGA STANDARD; PRT: 208 AA.  
ID FCP\_ISOGA  
AC 039709;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE FUCOXANTHIN-CHLOROPHYLL A-C BINDING PROTEIN, CHLOROPLAST PRECURSOR (FCP).  
GN FCP.  
OS Isochrysis galbana.  
OC Eukaryota: Haptophyceae; Isochrysidales; Isochrysis.  
OX NCBI\_TaxID=37099;  
RN [1]  
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
RC STRAIN-DUN;  
RX MEDLINE=94325461; PubMed=8049362;  
RA Laroche J., Henry D., Wyman K., Sukenik A., Falkowski P.;  
RT "Cloning and nucleotide sequence of a cDNA encoding a major fucoxanthin-chlorophyll a/c-containing protein from the chrysophyte  
RT Isochrysis galbana: implications for evolution of the cab gene family.";  
RT Plant Mol. Biol. 25:355-368(1994).  
CC -1- FUNCTION: THE LIGHT-HARVESTING COMPLEX (LHC) FUNCTIONS AS A LIGHT RECEPTOR. IT CAPTURES & DELIVERS EXCITATION ENERGY TO PHOTOSYSTEMS WITH WHICH IT IS CLOSELY ASSOCIATED. ENERGY IS TRANSFERRED FROM THE CAROTENOID AND CHL C (OR-B) TO CHL A AND THE PHOTOSYNTHETIC REACTION CENTERS WHERE IT IS USED TO SYNTHESIZE ATP AND REDUCING POWER.  
CC -1- SUBUNIT: THE LHC COMPLEX OF CHROMOPHYTIC ALGAE IS COMPOSED OF FUCOXANTHIN, CHLOROPHYLL A AND C BOUND NON-COVALENTLY BY PIGMENTS IN LHC; FUCOXANTHIN: CHLOROPHYLL C: CHLOROPHYLL A IS (0.6-1): (0.1-0.3): (1).  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE. FCPs ARE PROBABLY TRANSPORTED ACROSS THE ENDOPLASMIC RETICULUM MEMBRANES THAT SURROUND THE PLASTID VIA A SIGNAL PEPTIDE, FOLLOWED BY TRANSLOCATION ACROSS THE THYLAKOID MEMBRANE VIA A TRANSIT PEPTIDE.  
CC -1- INDUCTION: EXPRESSION IS INCREASED 5-FOLD UNDER CONDITIONS OF LOW LIGHT.  
CC -1- SIMILARITY: BELONGS TO THE FCP FAMILY OF LIGHT-HARVESTING PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X77333; CAAS4547.1; -  
DR InterPro: IPR001344; -  
DR Pfam: PF00504; chloroa\_b-bind. 2  
KW Light-harvesting polypeptide; Chloroplast; Photosynthesis;  
KW Photosystem II; Multigene family; Chlorophyll; Transmembrane;  
KW Thylakoid membrane; Transit peptide.  
FT TRANSIT 1 31 CHLOROPLAST (PROBABLE).  
FT CHAIN 32 208 FUCOXANTHIN-CHLOROPHYLL A-C BINDING PROTEIN.  
FT TRANSMEM 102 118 POTENTIAL.  
SQ SEQUENCE 208 AA; 22471 MM; 21A36700137A0F1B CRC64;

Query Match 66.1%; Score 39; DB 1; Length 208;  
Best Local Similarity 83.3%; Pred. No. 14;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
2 TWGQYW 7  
|||||

Db 198 TWGTYW 203

RESULT 13  
YB52\_MYCPN STANDARD; PRT: 794 AA.  
ID YB52\_MYCPN  
AC P75034;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL LIPOPROTEIN MPN152 PRECURSOR (E07\_09F794).  
GN MPN152 OR MP002.  
OS Mycoplasma pneumoniae.  
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2104;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 29342 / M129;  
RX MEDLINE=97105885; PubMed=8948633;  
RA Himmelreich R., Hilbert H., Plagens H., Pitkl E., Li B.-C.,  
RA Herrmann R.;  
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.";  
RT Nucleic Acids Res. 24:4420-4449(1996).  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID ANCHOR (POTENTIAL).  
CC -1- SIMILARITY: BELONGS TO THE MG185 / MG260 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AE000001; AA95650.1; -  
KW Hypothetical protein; Lipoprotein; Membrane; Signal.  
FT SIGNAL 1 22 POTENTIAL.  
FT CHAIN 23 794 HYPOTHETICAL LIPOPROTEIN MPN152.  
FT LIPID 23 23 N-ACYL DIGLYCERIDE (POTENTIAL).  
SQ SEQUENCE 794 AA; 86871 MM; 0DA72D12807E53BE CRC64;

Query Match 66.1%; Score 39; DB 1; Length 794;  
Best Local Similarity 66.7%; Pred. No. 48;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Y 1 KTWGQYW 9  
Db 268 KTWGQYW 276

RESULT 14  
POL\_STVA3 STANDARD; PRT: 100 AA.  
ID POL\_STVA3  
AC P12501;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE POL. POLYPROTEIN (CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16); DE REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)) (FRAGMENT).  
GN POL.  
OS Simian immunodeficiency virus (AGM385 isolate) (SIV-AGM).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11729;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89012184; PubMed=3172340;  
RA Daniel M.D., Li Y., Naidu Y.M., Durda P.J., Schmidt D.K.,  
RA Troup C.D., Silva D.P., Mackey J.J., Kestler H.W., Sehgal P.K.,

RA King N.W., Ohta Y., Hayami M., Desrosiers R.C.;  
 RT "Simian immunodeficiency virus from African green monkeys.";  
 RL J. Virol. 62:4123-4128(1988).  
 CC -1- MISCELLANEOUS; THIS IS AN AFRICAN GREEN MONKEY ISOLATE.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
 CC KNOWN AS THE RETROPEPSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
 CC -----  
 DR EMBL: M21311; AAA47583.1; -  
 DR HSSP: P04585; IKLM.  
 DR HIV: M21311; POLSAGM38.  
 DR MEROPS: A02.003; -  
 DR Interpro: IPR001969; -  
 DR PROSITE: PS00141; ASP\_PROTEASE; PARTIAL.  
 DR AIDS: Polypeptide; Hydrolyase; Aspartyl protease; Endonuclease;  
 KW Nuclease; Transferase; RNA-directed DNA polymerase.  
 FT NON\_TER 1 100  
 FT NON\_TER 1 100  
 SQ SEQUENCE 100 AA; 12023 MW; FC11C66792D37E3F CRC64;

Query Match 64.4%; Score 38; DB 1; Length 100;  
 Best Local Similarity 71.4%; Pred. No. 10;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 WGYMOV 9  
 Db 26 WADYMOV 32

RESULT 15  
 PT2B\_ARATH STANDARD; PRT; 585 AA.  
 AC P46032;  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 40, Last annotation update)  
 DE PEPTIDE TRANSPORTER PTR2-B (HISTIDINE TRANSPORTING PROTEIN).  
 GN PTR2-B OR NTRI OR AT2G02040 OR F14H20.11.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids II;  
 CC Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. LANDSBERG-ERECTA;  
 RA Song W., Steiner H.-Y., Zhang L., Naider F., Stacey G.,  
 RA Becker J.M.;  
 RL Submitted (XX-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. C24;  
 RX MEDLINE=94307379; PubMed=8033999;  
 RA Frommer W.B., Hummel S., Reentsch D.;  
 RT "Cloning of an Arabidopsis histidine transporting protein related to  
 RT nitrate and peptide transporters.";  
 RL FEBS Lett. 347:185-189(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
 RA Cronin L.A., Shen W., Vanden S.E., Umayam L., Taiton L.J., Gill J.E.,

RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
 RA Copenhagen G.P., Preus D., Nierman W.C., White O., Eisen J.A.,  
 RA Salberg S.L., Fraser C.M., Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 402:761-768(1999).  
 CC -1- FUNCTION: PEPTIDE TRANSPORT. HIGH AFFINITY, LOW CAPACITY  
 CC TRANSPORTER. CAN ALSO TRANSPORT HISTIDINE.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE PTR2 FAMILY OF TRANSPORTERS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
 CC -----  
 DR EMBL: L39082; AAB00858.1; -  
 DR EMBL: X77503; CA54634.1; -  
 DR EMBL: AC006532; AAD20096.1; -  
 DR Interpro: IPR00109; -  
 DR Pfam: PP00854; PTR2; 1.  
 DR PROSITE: PS01022; PTR2\_1; 1.  
 DR PROSITE: PS01023; PTR2\_2; 1.  
 KW Peptide transport; Transport; Transmembrane.  
 FT TRANSMEM 91 111  
 FT TRANSMEM 116 136  
 FT TRANSMEM 154 174  
 FT TRANSMEM 200 220  
 FT TRANSMEM 228 248  
 FT TRANSMEM 351 371  
 FT TRANSMEM 387 407  
 FT TRANSMEM 431 451  
 FT TRANSMEM 472 492  
 FT TRANSMEM 511 531  
 FT TRANSMEM 556 576  
 FT TRANSMEM 585 644  
 FT CONFLICT 334 334  
 SQ SEQUENCE 585 AA; 64421 MW; C58F8194776E2D97 CRC64;

Query Match 64.4%; Score 38; DB 1; Length 585;  
 Best Local Similarity 57.1%; Pred. No. 52;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 WGYMOV 9  
 Db 110 WGRYWT 116

RESULT 16  
 YEJH\_ECOLI STANDARD; PRT; 586 AA.  
 ID YEJH\_ECOLI  
 AC P33919; P36926; P36927; P76449;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE HYPOTHETICAL 66.4 KDA PROTEIN IN RSUA-RPLY INTERGENIC REGION.  
 GN YEJH.  
 OS Escherichia coli.  
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Escherichia.  
 OX NCBI\_Taxid=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / BHB2600.  
 RA Richterich P., Iakey N., Gryan G., Jaehn L., Mintz L., Robison K.,  
 RA Church G.M.;  
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;

RX MEDLINE-97426617; PubMed-9278503;  
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.,  
 RT "The complete genome sequence of *Escherichia coli* K-12.";  
 RL Science 277:1453-1474(1997).  
 CC -1- SIMILARITY: SOME SIMILARITY TO A PHAGE PROTEIN AND RESTRICTION-  
 CC MODIFICATION SYSTEMS.  
 CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS  
 CC THAT PRODUCE THREE SEPARATE ORFS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U00008; AAA16381.1; ALT\_FRAME.  
 DR EMBL: AE000308; AAC75245.1; -.  
 DR EcoGene: EG12045; yeJH.  
 DR InterPro: IPR001410; -.  
 DR InterPro: IPR001650; -.  
 DR Pfam: PF00270; DEAD; 1.  
 DR Pfam: PF00271; helicase\_C; 1.  
 DR Hypothetical protein.  
 KW Hypothetical protein.  
 SQ SEQUENCE 586 AA; 66413 MW; 2D173250F83333DF CRC64;

Query Match 64.4%; Score 38; DB 1; Length 586;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GQYQV 9  
 Db 562 GQYQV 567

RESULT 17  
 YE68\_METJA STANDARD; PRT; 1009 AA.  
 AC 058863;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE HYPOHETICAL PROTEIN MJ1468.  
 GN MJ1468.  
 OS Methanococcus jannaschii.  
 OC Archaea: Euryarchaeota: Methanococcales: Methanococcaceae;  
 OC Methanococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 2661.  
 RX MEDLINE-96337999; PubMed-8688087;  
 RA Bolt C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,  
 RA Uitterlisch T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*  
 RT *jannaschii*.";  
 RL Science 273:1058-1073(1996).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- SIMILARITY: CONTAINS 5 PRD DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U67588; AAB99478.1; -.  
 DR TIGR: MJ1468; -.  
 DR InterPro: IPR000601; -.  
 DR Pfam: PF00801; PKD; 3.  
 DR PROSITE: PS50093; PKD; 5.  
 KW Hypothetical protein; Transmembrane; Repeat.  
 FT TRANSMEM 6 26 POTENTIAL.  
 FT TRANSMEM 985 1005  
 FT DOMAIN 213 247 PKD 1.  
 FT DOMAIN 436 503 PKD 2.  
 FT DOMAIN 724 806 PKD 4.  
 FT DOMAIN 822 886 PKD 3.  
 FT DOMAIN 925 962 PKD 5.  
 FT DOMAIN 293 298 POLY-ASN.  
 SQ SEQUENCE 1009 AA; 115119 MW; 13E9B4933EAB7972 CRC64;

Query Match 64.4%; Score 38; DB 1; Length 1009;  
 Best Local Similarity 66.7%; Pred. No. 86;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGYWQ 8  
 Db 322 WGYWQ 327

RESULT 18  
 POL\_STVAG STANDARD; PRT; 1046 AA.  
 ID POL\_STVAG  
 AC P27980;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE POL POLYPROTEIN (CONTAINS: PROTEASE (RETROPEPIN) (EC 3.4.23.16);  
 DE REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)).  
 GN POL.  
 OS Simian immunodeficiency virus (AGM3 isolate) (SIY-AGM).  
 OC Viruses: Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11730;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-90232731; PubMed-2158689;  
 RA Balier M., Garber C., Mueller C., Cichutek K., Kurth R.;  
 RT "Complete nucleotide sequence of a simian immunodeficiency virus from  
 RT African green monkeys: a novel type of intragroup divergence.";  
 RL Virology 176:216-221(1990).  
 CC -1- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE  
 CC DETERMINED.  
 CC -1- MISCELLANEOUS: THIS IS AN AFRICAN GREEN MONKEY ISOLATE.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
 CC KNOWN AS THE RETROPEPIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M30931; AAA91914.1; -.  
 DR HSSP: P03366; 1HMV.  
 DR MEROPS: A02.003; -.  
 DR InterPro: IPR000477; -.  
 DR InterPro: IPR001037; -.  
 DR InterPro: IPR001584; -.  
 DR InterPro: IPR001669; -.

DR	Pfam: PF00075; naseh; 1.
DR	pFam: PF00665; rve; 1.
DR	pFam: PF00077; tvp; 1.
DR	pFam: PF00078; tvl; 1.
DR	PROSITE: PS00141; ASP_PROTEASE; 1.
DR	PROSITE: PS00175; ASP_PROT_RETROV; 1.
KW	AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
KW	Nuclease; Transferrase; RNA-directed DNA polymerase.
FT	ACT SITE 113 113 BY SIMILARITY
SQ	SEQUENCE 1047 AA; 118871 MW; A38DDDA39F266BE5 CRC64;
Qy	3 WGYWQV 9 1 1111
Db	592 WADYQV 598
RESULT 20	
POL_STVAL	
ID POL_STVAL STANDARD; PRT; 1057 AA.	
AC 002836;	
DT 01-JUL-1993 (Rel. 26, Created)	
DT 01-JUL-1993 (Rel. 26, Last sequence update)	
DT 01-OCT-2000 (Rel. 40, Last annotation update)	
DE POL PROTEIN [CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16); POLYMERASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)].	
GN POL.	
OC Simian immunodeficiency virus (isolate AGM / clone GRI-1) (SIV-AGM).	
OS Viruses; Retroviral viruses; Retroviridae; Lentivirus.	
OX NCBI_TaxID=31684;	
RN 11	
RP SEQUENCE FROM N.A.	
RA MEDLINE-91220680; PubMed-2024476; Fomsgard A., Hirsch V.M., Allan J.S., Johnson P.R.; "A highly divergent proviral DNA clone of SIV from a distinct species of African green monkey." Virology 182:397-402(1991)..	
RL -1 PPM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE DETERMINED	
CC -1 MISCELLANEOUS: THIS IS AN AFRICAN GREEN MONKEY ISOLATE. -1 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO KNOWN AS THE RETROPEPSIN FAMILY.	
CC CC	
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
DR EMBL: M66437; AAA91923.2; -; EMBL: MS8410; CAB26041.1; -; HSSP: P03366; IHMV.	
DR MEROPS: A02_003; -;	
DR InterPro: IPR000477; -	
DR InterPro: IPR001037; -	
DR InterPro: IPR001584; -	
DR InterPro: IPR001969; -	
DR InterPro: IPR001995; -	
DR InterPro: IPR002156; -	
DR Pfam: PF02022; Integrase_2n; 1.	
DR Pfam: PF00552; Integrase; 1.	
DR Pfam: PF00075; naseh; 1.	
DR Pfam: PF00665; rve; 1.	
DR Pfam: PF00077; tvp; 1.	
DR Pfam: PF00078; tvl; 1.	
DR PROSITE: PS00141; ASP_PROTEASE; 1.	
DR PROSITE: PS00175; ASP_PROT_RETROV; 1.	



KW AIDS: Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;  
KW Nuclease; Transferase; RNA-directed DNA polymerase.  
FT ACT\_SITE 114 114 BY SIMILARITY.  
SQ SEQUENCE 1057 AA; 120566 MW; 0550E954011FAA2C CRC64;

Query Match 64.4%; Score 38; DB 1; Length 1057;  
Best Local Similarity 62.5%; Pred. No. 89;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KTWGQYWG 8  
111 :11  
Db 590 KTWMMWQ 597

RESULT 21  
ID POL\_SIVAT STANDARD; PRT; 1061 AA.  
AC P05895;  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE POL POLYPROTEIN [CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);  
DE REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)]  
GN POL.  
OS Simian immunodeficiency virus (TPO-1 isolate) (STV-AGM).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11731;

RA FUKASAWA M., MURA T., HASEGAWA A., MORIKAWA S., TSUJIMOTO H.,  
RA MIKI K., KITAMURA T., HAYAMI M.;  
RA "Sequence of simian immunodeficiency virus from African green monkey,  
RT a new member of the HIV/SIV group."  
RL Nature 333:457-461(1988).  
CC -1 PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE  
CC DETERMINED.  
CC -1 MISCELLANEOUS: THIS IS AN AFRICAN GREEN MONKEY ISOLATE.  
CC -1 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
CC KNOWN AS THE RETROPEPSIN FAMILY.  
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

DR EMBL: X07805; CAA30658.1; ALT\_SEQ.  
DR PIR: B30045; GNLTG4.  
DR HSSP: P03366; IHMV.  
DR HIV: X07805; POLSAGMTY.  
DR MEROPS: A02.003; -  
DR InterPro: IPR000477; -  
DR InterPro: IPR001037; -  
DR InterPro: IPR001584; -  
DR InterPro: IPR001969; -  
DR InterPro: IPR001995; -  
DR InterPro: IPR002156; -  
DR Pfam: PF02022; Integrase\_Zn; 1.  
DR Pfam: PF00552; Integrase; 1.  
DR Pfam: PF00075; rnaesh; 1.  
DR Pfam: PF00665; rve; 1.  
DR Pfam: PF00077; rvp; 1.  
DR Pfam: PF00078; rvt; 1.  
DR PROSITE: PS00141; ASP\_PROTEASE; 1.  
DR PROSITE: PS50175; ASP\_PROT\_RETROV; 1.  
KW AIDS: Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;  
KW Nuclease; Transferase; RNA-directed DNA polymerase.  
FT CHAIN 1 210 PROTEASE.  
FT ACT\_SITE 134 134 BY SIMILARITY.

SO SEQUENCE 1061 AA; 120612 MW; 13DDDA104CB432A4 CRC64;

Query Match 64.4%; Score 38; DB 1; Length 1061;  
Best Local Similarity 71.4%; Pred. No. 90;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 WQYQV 9  
1111 :1111  
Db 613 WADYQV 619

RESULT 22  
ID UBIA\_ECOLI STANDARD; PRT; 290 AA.  
AC P26601;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE 4-HYDROXYBENZOATE OCTAPRENYLTRANSFERASE (EC 2.5.1.-) (4-HB  
DE POLYPRENYLTRANSFERASE).  
GN UBIA OR CYR.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
OX

RA "Cloning and sequencing of Escherichia coli ublc and purification of  
RT chorismate lyase."  
RL J. Bacteriol. 174:5309-5316(1992).  
RN

RA "Cloning and sequencing of Escherichia coli ublc and purification of  
RT chorismate lyase."  
RL J. Bacteriol. 174:5309-5316(1992).  
RN

RA "Cloning and sequencing of Escherichia coli ublc and purification of  
RT chorismate lyase."  
RL J. Bacteriol. 174:5309-5316(1992).  
RN

RA "Cloning and sequencing of Escherichia coli ublc and purification of  
RT chorismate lyase."  
RL J. Bacteriol. 174:5309-5316(1992).  
RN

RA "Cloning and sequencing of Escherichia coli ublc and purification of  
RT chorismate lyase."  
RL J. Bacteriol. 174:5309-5316(1992).  
RN

RA "Cloning and sequencing of Escherichia coli ublc and purification of  
RT chorismate lyase."  
RL J. Bacteriol. 174:5309-5316(1992).  
RN

RL J. Gen. Microbiol. 139:1795-1805(1993).  
 RN [7]  
 RP CHARACTERIZATION.  
 RX MEDLINE=94207029; PubMed=8155731;  
 RA Melzer M., Heide L.;  
 RT "Characterization of polyphenyldiphosphate:4-hydroxybenzoate  
 RT polyphenyltransferase from *Escherichia coli*.";  
 RL Polchim. Biophys. Acta 1212:93-107(1994).  
 RN [8]  
 RP CHARACTERIZATION.  
 RX MEDLINE=95072311; PubMed=7765507;  
 RA Suzuki K., Ueda M., Yuasa M., Nakagawa T., Kawamukai M., Matsuda H.;  
 RT "Evidence that *Escherichia coli* ubia product is a functional homolog  
 RT of yeast COQ2, and the regulation of ubia gene expression.";  
 RL Biosci. Biotechnol. Biochem. 58:1814-1819(1994).  
 CC -1- FUNCTIONAL SYNTHESIS OF 3-OCTAPRENYL-4-HYDROXYBENZOATE.  
 CC -1- CATALYTIC ACTIVITY: 4-HYDROXYBENZOATE + FARNESYLFARNESYLGERANIOL  
 CC -3-OCTAPRENYL-4-HYDROXYBENZOATE.  
 CC -1- COFACTOR: REQUIRES MAGNESIUM.  
 CC -1- PATHWAY: SECOND STEP IN UBIQUINONE BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.  
 CC -1- SIMILARITY: BELONGS TO THE UBIA PRENYLTRANSFERASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: M93136; AAA24712.1; -  
 DR EMBL: M93413; AAA24717.1; -  
 DR EMBL: X66619; CAA47182.1; -  
 DR EMBL: X57434; CAA40682.1; -  
 DR EMBL: X69522; CAA49270.1; -  
 DR EMBL: U00006; AAC43134.1; -  
 DR EMBL: AE000477; AAC77010.1; -  
 DR EMBL: M96268; AAA17028.1; -  
 DR EMBL: X63407; CAA45003.1; -  
 DR PIR: B42956; B42956.  
 DR PIR: S24361; S24361.  
 DR PIR: S25661; S25661.  
 DR PIR: S31432; S31432.  
 DR PIR: JC2316; JC2316.  
 DR EcoGene: EG11370; ubia.  
 DR InterPro: IPR000537; -  
 DR Pfam: PF01040; Ubia; 1.  
 DR PROSITE: PS00943; Ubia; 1.  
 KW Ubiquinone biosynthesis; Transferase; Transmembrane; Inner membrane;  
 KW Magnesium.  
 FT TRANSMEM 23 43 POTENTIAL.  
 FT TRANSMEM 46 66 POTENTIAL.  
 FT TRANSMEM 100 120 POTENTIAL.  
 FT TRANSMEM 141 161 POTENTIAL.  
 FT TRANSMEM 163 183 POTENTIAL.  
 FT TRANSMEM 213 233 POTENTIAL.  
 FT TRANSMEM 234 254 POTENTIAL.  
 FT TRANSMEM 268 288 POTENTIAL.  
 SQ SEQUENCE 290 AA; 32511 MW; F10PEDID7A30E115 CRC64;

Query Match 62.7%; Score 37; DB 1; Length 290;  
 Best Local Similarity 57.1%; Pred. No. 39;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 WGYWYOV 9  
 DB 237 WGYWYSI 243

RESULT 23  
 SLEB\_BACSU

ID SLEB\_BACSU STANDARD; PRT; 305 AA.  
 AC P50739;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE SPORE-CORTEX-LYTIC ENZYME PRECURSOR.  
 GN SLEB.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=96427343; PubMed=8830707;  
 RA Moriyama R., Hattori A., Miyata S., Kudoh S., Makino S.;  
 RT "A gene (sleb) encoding a spore cortex-lytic enzyme from *Bacillus*  
 RT subtilis and response of the enzyme to L-alanine-mediated  
 RT germination.";  
 RL J. Bacteriol. 178:6059-6063(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / MARBURG;  
 RX MEDLINE=96349105; PubMed=8760912;  
 RA Sorokin A.V., Azevedo V., Zumbstein E., Galleron N., Ehrlich S.D.,  
 RA Serror P.;  
 RT "Sequence analysis of the *Bacillus subtilis* chromosome region between  
 RT the *serA* and *kdg* loci cloned in a yeast artificial chromosome.";  
 RL Microbiology 142:2005-2016(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: D79978; BAA11473.1; -  
 DR EMBL: L47648; AAC83957.1; -  
 DR EMBL: Z99115; CAB14209.1; -  
 DR EMBL: Z99116; CAB14225.1; -  
 DR Subtilist; BG11439; sleb.  
 DR InterPro: IPR002477; -  
 DR Pfam: PF01471; PG\_binding\_1; 1.  
 KW Signal.  
 FT SIGNAL 1 29 POTENTIAL.  
 FT CHAIN 30 305 SPORE-CORTEX-LYTIC ENZYME.  
 SQ SEQUENCE 305 AA; 34001 MW; 9DF1305975F5BE16 CRC64;

Query Match 62.7%; Score 37; DB 1; Length 305;  
 Best Local Similarity 71.4%; Pred. No. 41;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 WGYWYOV 9  
 DB 67 WGYWYAV 73

RESULT 24  
 LIP3\_DROME STANDARD; PRT; 394 AA.  
 AC 046108;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE LIPASE 3 PRECURSOR (EC 3.1.1.-).  
 GN LIP3 OR CG8823.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=CANTON-S;  
 RX MEDLINE=98227313; PubMed=9566193;  
 RA Pistillo D., Manzi A., Tino A., Pilo Boyl P., Graziani F., Malva C.;  
 RT "The Drosophila melanogaster lipase homologs: a gene family with  
 RT tissue and developmental specific expression.";  
 RL J. Mol. Biol. 276:877-885(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brannon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.T., Beron P.V., Bernier B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jallat M., Kalush F., Karpen G.H., Ke Z., Kenton J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrstkas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 CC -1- TISSUE SPECIFICITY: FAT BODY.  
 CC -1- DEVELOPMENTAL STAGE: ONLY AT LARVAL STAGES.  
 CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,  
 CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO  
 CC LIPIDOPROTEIN EGG-SPECIFIC AND YOLK PROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: Y14367; CAAT74737.1;  
 CC EMBL: AE003699; AAF54935.1;  
 CC FLYBASE: FBgn003495; Lip3.  
 CC InterPro: IPR000073;  
 CC InterPro: IPR000379;  
 CC InterPro: IPR000734;  
 CC Pfam: PF00561; abhydrolase\_1.  
 CC PROSITE: PS00120; LIPASE\_SER; 1.

KW Hydrolase: Lipid degradation; Signal: Glycoprotein.  
 FT SIGNAL 1 20  
 FT CHAIN 21 394  
 FT ACT\_SITE 164 164  
 FT ACT\_SITE 369 369  
 FT CARBOHYD 131 131  
 SQ SEQUENCE 394 AA; 44901 MW; A718D1D743673802 CRC64;  
  
 Query Match 62.7%; Score 37; DB 1; Length 394;  
 Best Local Similarity 55.6%; Pred. No. 52;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
 QY 1 KTWGQYMQ 9  
 DB 120 KTWGQYMQ 128  
  
 RESULT 25  
 YFE9\_YEAST STANDARD; PRT; 623 AA.  
 AC P43534;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE HYPOTHETICAL 70.3 KDA PROTEIN IN ALR2-EMP47 INTERGENIC REGION.  
 GN YF049W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=95400292; PubMed=7670463;  
 RA Murakami Y., Natou M., Hagihara H., Shibata T., Ozawa M.,  
 RA Sasakura S.-I., Sasakura M., Tsuchiya Y., Soeda E., Yokoyama K.,  
 RA Yamazaki M., Tashiro H., Eki T.;  
 RT "Analysis of the nucleotide sequence of chromosome VI from  
 RT Saccharomyces cerevisiae.";  
 RL Nat. Genet. 10:261-268(1995).  
 CC -1- SIMILARITY: TO YEAST NPL6.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: D50617; BAA09192.1;  
 CC SGD: S0001845; YF049W.  
 KW Hypothetical protein.  
 SQ SEQUENCE 623 AA; 70275 MW; ABAC00CAC8F0ED0E CRC64;  
  
 Query Match 62.7%; Score 37; DB 1; Length 623;  
 Best Local Similarity 62.5%; Pred. No. 78;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 1 KTWGQYMQ 8  
 DB 491 KTWGQYMQ 498  
  
 RESULT 26  
 YD64\_MYCPN STANDARD; PRT; 677 AA.  
 AC P75417;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE HYPOTHETICAL PROTEIN MPN364 (H91\_ORF677).  
GN MPN364 OR MP472.  
OS Mycoplasma pneumoniae.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2104;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 29342 / M129;  
RX MEDLINE=97105885; PubMed=8948633;  
RA Himmelfrich R., Hilbert H., Plagens H., Pirk E., Li B.-C.,  
Herrmann R.,  
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
pneumoniae.";  
RL Nucleic Acids Res. 24:4420-4449(1996).  
CC -1- SIMILARITY: BELONGS TO THE MG185 / MG260 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AE000047; AAB96120.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 677 AA; 75591 MW; E785B68BD679F04D CRC64;  
  
Query Match 62.7%; Score 37; DB 1; Length 677;  
Best Local Similarity 66.7%; Pred. No. 85;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 KTWGQYMOV 9  
DB 153 KTWGQYMOV 161  
|||||  
POL\_HV2CA STANDARD; PRT; 1034 AA.  
ID POL\_HV2CA  
AC P24107;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE POL. POLYPROTEIN [CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);  
DE REVERSE TRANSCRIPTIONASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)].  
GN POL.  
OS Human immunodeficiency virus type 2 (isolate CAM2) (HIV-2).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11715;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91170959; PubMed=2005437;  
RA Tristem M., Hill F., Karpas A.,  
RT "Nucleotide sequence of a Guinea-Bissau-derived human  
immunodeficiency virus type 2 proviral clone (HIV-2CAM2).";  
RL J. Gen. Virol. 72:721-724(1991).  
CC -1- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE  
CC DETERMINED.  
CC -----  
CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR TGA MAY OCCUR BETWEEN  
CC 564-TLE AND 565-GLY.  
CC -----  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
CC KNOWN AS THE RETROPEPSIN FAMILY.  
CC -----  
DR PIR: B38475; GNLJCA.  
DR HSSP: P04584; GNLJCA.  
DR MEROPS: A02.002; -.  
DR InterPro: IPR000477; -.  
DR InterPro: IPR001037; -.  
DR InterPro: IPR001584; -.  
DR InterPro: IPR001969; -.  
DR InterPro: IPR001969; -.  
DR InterPro: IPR001969; -.  
DR InterPro: IPR001969; -.

DR InterPro: IPR002156; -.  
DR Pfam: PF02022; Integrase; Zn. 1.  
DR Pfam: PF00552; Integrase; 1.  
DR Pfam: PF00075; rnsesh; 1.  
DR Pfam: PF00665; rve; 1.  
DR Pfam: PF00077; rvp; 1.  
DR Pfam: PF00078; rvc; 1.  
DR Pfam: PF00078; rvc; 1.  
DR PROSITE: PS00141; ASP\_PROTEASE; 1.  
DR PROSITE: PS00175; ASP\_PROT\_RETROV; 1.  
DR AIDS: Polypeptide; Hydrolyase; Aspartyl protease; Endonuclease;  
KW Nuclease; Transferase; RNA-directed DNA polymerase.  
FT CHAIN 85  
FT ACT\_SITE 109 109  
SQ SEQUENCE 1034 AA; 117195 MW; 3514E566AA6D7C86 CRC64;  
  
Query Match 62.7%; Score 37; DB 1; Length 1034;  
Best Local Similarity 71.4%; Pred. No. 12e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 3 MGQYMOV 9  
DB 583 MGQYMOV 589  
|||||  
POL\_HV2KR STANDARD; PRT; 1035 AA.  
ID POL\_HV2KR  
AC 074120;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE POL. POLYPROTEIN [CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);  
DE REVERSE TRANSCRIPTIONASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)].  
GN POL.  
OS Human immunodeficiency virus type 2 (isolate KR) (HIV-2).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=73484;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kraus G.K., Talbot R., Leavitt M., Luzznick L., Schmidt A.,  
RA Badel P., Bartz C., Morton W., Wong-Staal F., Looney D.J.;  
RL Submitted (APR-1995) to the EMBL/GenBank/DBD databases.  
CC -1- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE  
CC DETERMINED.  
CC -----  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
CC KNOWN AS THE RETROPEPSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U20447; AAA64576.1; -.  
DR HSSP: P04584; 1JLD.  
DR InterPro: IPR000477; -.  
DR InterPro: IPR001037; -.  
DR InterPro: IPR001584; -.  
DR InterPro: IPR001969; -.  
DR InterPro: IPR001969; -.  
DR InterPro: IPR001969; -.  
DR Pfam: PF02022; Integrase; Zn. 1.  
DR Pfam: PF00552; Integrase; 1.  
DR Pfam: PF00075; rnsesh; 1.  
DR Pfam: PF00075; rnsesh; 1.  
DR Pfam: PF00665; rve; 1.  
DR Pfam: PF00077; rvp; 1.  
DR Pfam: PF00078; rvc; 1.  
DR PROSITE: PS00141; ASP\_PROTEASE; 1.  
DR PROSITE: PS00175; ASP\_PROT\_RETROV; 1.  
DR AIDS: Polypeptide; Hydrolyase; Aspartyl protease; Endonuclease;  
KW AIDS; Polypeptide; Hydrolyase; Aspartyl protease; Endonuclease;

KW Nuclease; Transferase; RNA-directed DNA polymerase.  
 FT CHAIN 85 183 PROTEASE.  
 FT ACT SITE 109 109 BY SIMILARITY  
 SQ SEQUENCE 1035 AA; 117632 MW; 696E6DC6CAFB06CF CRC64;

Query Match 62.7%; Score 37; DB 1; Length 1035;  
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 WGOYMOV 9  
 Db 584 WDNWMOV 590

RESULT 29  
 ID POL\_HV2NZ STANDARD; PRT; 1035 AA.  
 AC P05962; Q85571;  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE POL POLYPROTEIN [CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);  
 DE REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)]  
 GN POL.  
 OS Human immunodeficiency virus type 2 (isolate NIH-2) (HIV-2).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11719;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88320359; PubMed=1261862;  
 RA Zagury J.F., Franchini G., Reitz M.S. Jr., Collalti E., Starch B.R.,  
 RA Hall L., Fargnoli K., Jagodzinski L.L., Guo H.-G., Laure F.,  
 RA Ayra S.K., Josephs S.F., Zagury D., Wong-Staal F., Gallo R.C.;  
 RT "Genetic variability between isolates of human immunodeficiency virus  
 RT Proc. Natl. Acad. Sci. U.S.A. 85:5941-5945(1988).  
 RL -1- PWM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE  
 CC DETERMINED.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
 CC KNOWN AS THE RETROPEPSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J03654; AAB00755.1; ALT\_INT.  
 DR HSSP: P04584; 1JLD.  
 DR HIV: J03654; POLS2NINZ.  
 DR MEROPS: A02.002; -  
 DR InterPro: IPR000477; -  
 DR InterPro: IPR001037; -  
 DR InterPro: IPR001584; -  
 DR InterPro: IPR001969; -  
 DR InterPro: IPR001995; -  
 DR InterPro: IPR002156; -  
 DR Pfam: PF02022; Integrase\_2n; 1.  
 DR Pfam: PF00052; Integrase; 1.  
 DR Pfam: PF00075; rna5h; 1.  
 DR Pfam: PF00065; rve; 1.  
 DR Pfam: PF00077; rvp; 1.  
 DR Pfam: PF00078; rvt; 1.  
 DR PROSITE: PS00141; ASP\_PROTASE; FALSE\_NEG.  
 DR PROSITE: PS00175; ASP\_PROT\_RETROV; 1.  
 KW AIDS: Polypeptide; Hydrolase; Aspartyl protease; Endonuclease;  
 KW Nuclease; Transferase; RNA-directed DNA polymerase.  
 FT CHAIN 85 183 PROTEASE.  
 FT ACT\_SITE 109 109 BY SIMILARITY.  
 SQ SEQUENCE 1035 AA; 117323 MW; E3DC4E2DF457F6BA CRC64;

Query Match 62.7%; Score 37; DB 1; Length 1035;  
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 WGOYMOV 9  
 Db 584 WDNWMOV 590

RESULT 30  
 ID POL\_HV2SB STANDARD; PRT; 1035 AA.  
 AC P12451; Q85570;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE POL POLYPROTEIN [CONTAINS: PROTEASE (RETROPEPSIN) (EC 3.4.23.16);  
 DE REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)]  
 GN POL.  
 OS Human immunodeficiency virus type 2 (isolate SBLISY) (HIV-2).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11718;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89184641; PubMed=2648404;  
 RA Franchini G., Fargnoli K.A., Giombini F., Jagodzinski L., de Rossi A.,  
 RA Bosch M., Alberfeld G., Fenyo A.M., Albert J., Gallo R.C.,  
 RA Wong-Staal F.;  
 RT "Molecular and biological characterization of a replication competent  
 RT human immunodeficiency type 2 (HIV-2) proviral clone."  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:2433-2437(1989).  
 RL -1- PWM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE  
 CC DETERMINED.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
 CC KNOWN AS THE RETROPEPSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J04498; AAB00746.1; ALT\_INT.  
 DR HSSP: P04584; 1JLD.  
 DR HIV: J04498; POLS2ISY.  
 DR MEROPS: A02.002; -  
 DR InterPro: IPR000477; -  
 DR InterPro: IPR001037; -  
 DR InterPro: IPR001584; -  
 DR InterPro: IPR001969; -  
 DR InterPro: IPR001995; -  
 DR InterPro: IPR002156; -  
 DR Pfam: PF02022; Integrase\_2n; 1.  
 DR Pfam: PF00052; Integrase; 1.  
 DR Pfam: PF00075; rna5h; 1.  
 DR Pfam: PF00065; rve; 1.  
 DR Pfam: PF00077; rvp; 1.  
 DR Pfam: PF00078; rvt; 1.  
 DR PROSITE: PS00141; ASP\_PROTASE; 1.  
 DR PROSITE: PS00175; ASP\_PROT\_RETROV; 1.  
 KW AIDS: Polypeptide; Hydrolase; Aspartyl protease; Endonuclease;  
 KW Nuclease; Transferase; RNA-directed DNA polymerase.  
 FT CHAIN 85 183 PROTEASE.  
 FT ACT\_SITE 109 109 BY SIMILARITY.  
 SQ SEQUENCE 1035 AA; 117552 MW; 9DA9BD9DBE4393 CRC64;

Query Match 62.7%; Score 37; DB 1; Length 1035;  
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;

Matches	5;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	3	WGOYWOY	9						
Db	584	WDNYWOY	590						

Search completed: July 3, 2001, 10:28:38  
Job time: 323 sec

7

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:23:58 ; Search time 33.34 Seconds  
(without alignments)  
16.365 Million cell updates/sec

Title: US-09-214-836-9

Perfect score: 59

Sequence: 1 KTWGQYQGV 9

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq.0601.\*  
1: /SIDSR/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDSR/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDSR/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDSR/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDSR/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDSR/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDSR/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDSR/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDSR/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDSR/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDSR/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDSR/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDSR/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDSR/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDSR/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDSR/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDSR/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDSR/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDSR/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDSR/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDSR/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDSR/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	9	16 AAR78644	Immunogenic peptid
2	59	100.0	9	19 AAM77119	gp100/Pmel17 synth
3	59	100.0	9	19 AAM78850	PMEL 17 (gp100) pr
4	59	100.0	9	19 AAM70010	Melanoma-associate
5	59	100.0	9	19 AAM54598	Peptide 3 from gp
6	59	100.0	9	19 AAM54570	Melanoma associate
7	59	100.0	9	20 AAY49663	Tumour antigenic p
8	59	100.0	9	20 AAY53524	Human melanoma pme
9	59	100.0	9	20 AAY47616	Immunogenic peptid
10	59	100.0	9	20 AAY33172	Human gp100-Pmel11
11	59	100.0	9	20 AAY40211	Amino acid sequenc

12	59	100.0	9	20 AAY26867	Melanoma-derived 1
13	59	100.0	9	20 AAY01753	Exemplary antigen
14	59	100.0	9	20 AAY00715	Tumour antigen boe
15	59	100.0	9	20 AAY10449	HLA Class I motif
16	59	100.0	9	21 AAB33662	MHC class I associ
17	59	100.0	9	21 AAB33679	Cytotoxic T lympho
18	59	100.0	9	21 AAB08694	Antigenic peptide
19	59	100.0	9	21 AAY71520	Human gp100 Pmel11
20	59	100.0	9	21 AAB02622	Tumour associated
21	59	100.0	9	21 AAY90803	Human leukocyte an
22	59	100.0	9	21 AAY92299	gp100-Pmel117 anti
23	59	100.0	9	21 AAY84296	Tumour associated
24	59	100.0	9	21 AAY82979	gp100(Pmel117) tum
25	59	100.0	9	21 AAY56661	gp100-Pmel117 gen
26	59	100.0	9	22 AAB31354	Exemplary antigen
27	59	100.0	10	16 AAR78643	Immunogenic peptid
28	59	100.0	10	16 AAR84209	gp100 melanoma ant
29	59	100.0	11	16 AAR78642	Immunogenic peptid
30	59	100.0	30	22 AAB61647	gp100 peptide #1.
31	59	100.0	661	16 AAR84855	MART-1 melanoma an
32	59	100.0	661	16 AAR78646	Melanoma associate
33	59	100.0	661	20 AAY31977	Human melanoma ant
34	59	100.0	662	20 AAY42627	Human melanoma ant
35	59	100.0	668	18 AAM38164	Pmel17 encoded by
36	55	93.2	9	19 AAM42542	Gp 100 epitope (re
37	55	93.2	625	20 AAY31978	Mouse melanoma ant
38	55	93.2	661	16 AAR84854	MART-1 melanoma an
39	54	91.5	8	16 AAR78641	Immunogenic peptid
40	54	91.5	9	16 AAR78645	Immunogenic peptid
41	54	91.5	9	16 AAR84804	Modified MART-1 me
42	54	91.5	9	16 AAR84805	Modified MART-1 me
43	54	91.5	9	16 AAR84806	Modified MART-1 me
44	54	91.5	9	16 AAR84807	Modified MART-1 me
45	54	91.5	9	16 AAR84808	Modified MART-1 me

#### ALIGNMENTS

RESULT 1	
AAAR78644	
ID AAR78644 standard; Protein; 9 AA.	
XX	
AC AAR78644;	
XX	
22-JAN-1996 (first entry)	
XX	
DE Immunogenic peptide of melanoma associated antigen gp100.	
XX	
KW Melanoma: antigen; vaccine; immunogen; primer; probe; detection;	
KW Identification; tumour; gp100.	
XX	
OS Homo sapiens.	
XX	
PN EPE68350-A1.	
XX	
23-AUG-1995.	
XX	
14-FEB-1995; 95EP-0200348.	
XX	
21-DEC-1994; 94EP-0203709.	
XX	
PR 16-FEB-1994; 94EP-0200337.	
XX	
PA (ALKU ) AKZO NOBEL NV.	
XX	
PI Adema GJ, Figdor CG;	
XX	
DR WPI: 1995-284790/38.	
XX	
N-PSDB: AAO96055.	
XX	
Melanoma associated antigen gp100 - used in vaccines and for the	
PT detection of tumours	
XX	

PS Claim 5; Page 31; 40pp; English.

XX Immunogenic peptides derived from the melanoma associated antigen  
CC (See AAR78639-45) may be used in the production of vaccines.  
CC Nucleotide sequences encoding the immunogenic peptides may be used  
CC as primers and probes in the detection of melanoma cells. Tumour  
CC infiltrating lymphocytes capable of binding to the melanoma  
CC associated antigen can be cultured ex vivo and returned to melanoma  
CC particles, and when radiolabelled, they may be used to identify  
CC tumour deposits.

XX  
SQ Sequence 9 AA;

Query Match 100.0%; Score 59; DB 16; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGOTQV 9  
111111111  
DB 1 ktwgqywyv 9

## RESULT 2

AAW77119  
ID AAW77119 standard; peptide; 9 AA.

XX  
AC AAW77119;

XX 16-NOV-1998 (first entry)

DE gp100/Pmel17 synthetic peptide epitope 1.

XX Tyrosinase; tyrosinase cytotoxic lymphocyte response;

KW cytotoxic T lymphocyte; cysteine-depleted; melanoma.

XX Synthetic.

OS  
PN WO9833810-A2.

XX 06-AUG-1998.

XX 29-JAN-1998; 98WO-US01592.

XX 30-JAN-1997; 97US-0037781.

XX (UYVI-) UNIV VIRGINIA PATENT FOUND.

XX Engelhard VH, Hunt DF, Kittlesen D, Slingluff CL;

XX WPI; 1998-437388/37.

DR  
XX Disease specific immunogen - comprises disease specific cytotoxic T  
PT lymphocyte epitope used to elicit melanoma specific CTL response

XX  
PS Disclosure; Page 27; 93pp; English.

XX The peptide epitope AAW77119-W77138 were created for human  
CC tumour-specific cytotoxic T lymphocyte response. These peptides are  
CC cysteine- depleted mutants of a native disease-specific CTL epitope. The  
CC response than the native epitope. The epitopes can be used in a  
CC disease-specific immunogen to protect a mammal against disease in  
CC particular melanomas. The peptides may also be used to screen a sample  
CC for the presence of an antigen with the same epitope, or with a different  
CC cross-reactive epitope.

XX  
SQ Sequence 9 AA;

Query Match 100.0%; Score 59; DB 19; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGOTQV 9  
111111111  
DB 1 ktwgqywyv 9

## RESULT 3

AAW78850  
ID AAW78850 standard; peptide; 9 AA.

XX  
AC AAW78850;

XX 17-NOV-1998 (first entry)

DE PMEL 17 (gp100) protein fragment 154-162.

KW Microparticle; delivery; polymeric matrix; autoantigen; tumour antigen;  
KW class II associated peptide; pathogen; gene therapy; genetic disease;  
KW infection; downregulation; immune response.

XX Homo sapiens.

OS Synthetic.

PN WO9831398-A1.

XX 23-JUL-1998.

XX 22-JAN-1998; 98WO-US01499.

XX 06-JAN-1998; 98US-0003253.

XX 22-JAN-1997; 97US-0787547.

XX (PANG-) PANGAEA PHARM INC.

XX Curley JM, Hedley ML, Langer RS, Lunsford LB;

XX WPI; 1998-427556/36.

DR  
XX New preparations of microparticles - comprising a synthetic polymer  
PT matrix and nucleic acid comprising an expression vector for use in  
PT gene therapy

XX  
PS Disclosure; Page 10; 101pp; English.

XX A microparticle preparation (MP) has been developed, consisting of  
CC microparticles having a diameter of less than 100 mu m. The MP  
CC comprises: (a) a polymeric matrix (PM) consisting of one or more  
CC synthetic polymers having a solubility in water of less than 1 mg/L; and  
CC (b) an expression vector selected from RNA molecules (at least 50% of  
CC which are closed circles) or circular plasmid DNA (at least 50% of which  
CC are supercoiled). Also described is a MP of at most 20 microns in  
CC diameter, comprising: (a) a PM; and (b) a NAM comprising an expression  
CC control sequence operatively linked to a coding sequence, where the  
CC coding sequence encodes an expression product selected from: (1) a  
CC polypeptide at least 7 amino acids in length, having a sequence identical  
CC to the sequence of: (1) a fragment of a naturally-occurring mammalian  
CC protein; or (11) a fragment of a naturally-occurring protein from an  
CC infectious agent which infects a mammal; (2) a peptide having a length  
CC and sequence which permits it to bind to an MHC class I or II molecule;  
CC and (3) the polypeptide or the peptide linked to a trafficking sequence.  
CC AAW69763 to AAW69765, and AAW78793 to AAW78897 are peptide fragments for  
CC use in the present invention. The MPs are highly effective vehicles for  
CC the delivery of polynucleotides into phagocytic cells. They can be used  
CC for gene therapy, e.g. for treating genetic diseases, infections or  
CC tumours or for downregulating an immune response.

XX  
SQ Sequence 9 AA;

Query Match 100.0%; Score 59; DB 19; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 KTWGQYMOV 9  
 |||||  
 Db 1 ktwgqyqv 9

## RESULT 4

AAW70010  
 ID AAW70010 standard; peptide: 9 AA.

AC AAW70010;

DT 22-OCT-1998 (first entry)

DE Melanoma-associated antigen gp100 derived HLA-A2.1 binding peptide 1.

XX Cytotoxic T lymphocyte; CTL; major histocompatibility complex; MHC;  
 KW human leukocyte antigen; HLA; tumour associated antigen; cancer;  
 KW antigen presenting cell; APC; immunogenic peptide; immune disorder;  
 KW viral infection; AIDS; hepatitis; bacterial infection; malaria;  
 KW fungal infection; tuberculosis; melanoma; gp100.

OS Synthetic.  
 OS Homo sapiens.

PN W09833888-A1.

PD 06-AUG-1998.

PF 30-JAN-1998; 98WO-US01959.

PR 31-JAN-1997; 97US-0036696.

XX (EPIM-) EPIMMUNE INC.

PI Celis E, Sette A, Sidney J, Southwood S, Tsai Y;

DR WPI: 1998-437445/37.

XX Production of antigen-specific cytotoxic T cells - by incubating  
 PT immunogenic peptide(s) from antigen that binds class I major  
 PT histocompatibility complex molecules with pre-treated antigen  
 PT presenting cells

PS Example 4; Page 62; 104pp: English.

XX Sequences shown in AAW70010 to AAW70026 represent peptides derived from  
 CC melanoma-associated antigen gp100 that can bind to a human leukocyte  
 CC antigen (HLA), HLA-A2.1. The peptides are used to exemplify the method  
 CC of invention of producing antigen-specific cytotoxic T cells (CTLs) in  
 CC vitro. The method comprises contacting immunogenic peptides from an  
 CC antigen that binds class I major histocompatibility complex (MHC)  
 CC molecules with antigen presenting cells (APCs) pretreated with  
 CC pretreatment growth factors, and incubating the APCs with purified CD8  
 CC cells in the presence of at least 2 incubation growth factors, thereby  
 CC producing antigen-specific CTLs. A method for specifically killing  
 CC target cells in a human patient is also provided which comprises  
 CC obtaining a fluid sample containing CTLs from a patient, contacting the  
 CC cytotoxic T cells with APCs pretreated with pre-treatment growth  
 CC factors, where the APCs comprise class I MHC molecules. The pretreated  
 CC APCs are incubated with the cytotoxic growth factors, thereby producing  
 CC activated CTLs which are contacted with a carrier to form a composition.  
 CC The composition can then be administered to the patient. The activated  
 CC CTLs can be used for treating cancers, immune disorders, viral  
 CC infections, AIDS, hepatitis, bacterial infection, fungal infection,  
 CC malaria or tuberculosis.

XX Sequence 9 AA;

Query Match 100.0%; Score 59; DB 19; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGQYMOV 9  
 |||||  
 Db 1 ktwgqyqv 9

## RESULT 5

AAW54598  
 ID AAW54598 standard; peptide: 9 AA.

AC AAW54598;

DT 25-SEP-1998 (first entry)

DE Peptide 3 from gp 100/Pmel-17.

XX Mannose; antigen; antigen-presenting cell; mannosylated peptide; T cell;  
 KW vaccine; treatment.

OS Synthetic.

PN W09813378-A1.

PD 02-APR-1998.

PF 25-SEP-1997; 97WO-NL00536.

PR 26-SEP-1996; 96EP-0202701.

XX (UYLE-) RIJKSUNIV LEIDEN.

PI Drifhout JW, Konig F;

DR WPI: 1998-230631/20.

XX Increasing uptake and presentation of antigen(s) - by adding mannose  
 PT residue(s) to antigen for increasing T cell response, useful in,  
 PT e.g. vaccines against viral infection(s)

PS Disclosure; Page 24; 47pp: English.

XX The peptides AAW5459-W54809 are examples of peptides to which at least  
 CC 1 (preferably 2) mannose can be attached to increase their uptake as  
 CC antigens by antigen-presenting cells. Uptake of agonist mannosylated  
 CC peptides will increase the T cell response, whereas uptake of antagonist  
 CC peptides blocks the T cell response. Blocking binding of immunogenic  
 CC autoantigens can be used in treatment of type I diabetes, rheumatoid  
 CC arthritis, graft rejection etc., also to induce T-cell non-  
 CC responsiveness. Vaccines containing mannosylated antigen are used to  
 CC prevent or treat infections by, e.g. bacteria, viruses, fungi, helminths  
 CC and parasites.

XX Sequence 9 AA;

Query Match 100.0%; Score 59; DB 19; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGQYMOV 9  
 |||||  
 Db 1 ktwgqyqv 9

## RESULT 6

AAW5770  
 ID AAW5770 standard; peptide: 9 AA.

AC AAW5770;

DT 22-JUN-1998 (first entry)

DE Melanoma associated peptide analogue #1.

KW Metastatic melanoma; peptide analogue; vaccine; cancer; diagnosis;  
KM antigen; CTL; immunogenic; viral disease.  
XX  
OS Synthetic.  
XX Homo sapiens.  
PN WO9802538-A1.  
XX  
PD 22-JAN-1998.  
XX  
PF 08-JUL-1997; 97WO-EP03712.  
XX  
PR 11-JUL-1996; 96EP-0201945.  
XX  
PA (ALKU ) AKZO NOBEL NV.  
XX  
PI Adema GJ, Figdor CG;  
XX  
DR WPI; 1998-110586/10.  
XX  
PT Melanoma associated peptide analogues - useful in vaccines against  
PT melanoma  
XX  
PS Claim 1; Figure 1; 47pp; English.  
XX  
CC This sequence represents a specifically claimed example of a novel  
CC peptide, which is immunogenic with lymphocytes directed against  
CC metastatic melanomas. It is characterised in that it comprises at least  
CC a part of the following sequence, where the amino acid at position 2 or 8  
CC is substituted: Lys-Thr-Trip-Gly-Gln-Tyr-Trip-Gln-Val. Vaccines comprising  
CC the peptide, an epitope of the peptide, nucleotide sequence encoding the  
CC peptide, or an antigen presenting cell preloaded with the peptide or  
CC antibody, as above, are useful for cancer, particularly melanoma,  
CC treatment. The peptides can also be used to generate antigen reactive  
CC tumour infiltrating lymphocytes, which can also be used in vaccines. The  
CC peptides can be exploited to elicit native epitope reactive CTL. Usage  
CC of the peptides with improved immunogenicity may contribute to the  
CC development of CTL-epitope based vaccines in viral disease and cancer.  
CC  
SQ Sequence 9 AA;  
  
Query Match 100.0%; Score 59; DB 19; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KTWGQYQYQV 9  
DB 1 ktwgqyqwv 9  
  
RESULT 7  
AAV49663  
ID AAV49663 standard; peptide; 9 AA.  
XX  
AC AAV49663;  
XX  
DT 14-JAN-2000 (first entry)  
XX  
DE Tumour antigenic peptide SEQ ID NO:30.  
XX  
KM Human; sdph3.10; SAGE; sdph3.8; HAGE; sdph3.5; TRAP; sarcoma;  
KM tumour rejection antigen precursor; tumour associated nucleic acid;  
KM carcinoma; cancer; immune response; diagnosis.  
XX  
OS Synthetic.  
XX  
PN WO9953061-A2.  
XX  
PD 21-OCT-1999.  
XX  
PF 14-APR-1999; 99WO-US08163.  
XX

PR 15-APR-1998; 98US-0060706.  
PR 27-JUL-1998; 98US-0122989.  
PR 30-OCT-1998; 98US-0183706.  
PR 30-OCT-1998; 98US-0183789.  
XX  
PA (LUDW-) LUDWIG INST CANCER RES.  
XX  
PI Martelange V, De Smet C, Boon-Falleur T;  
XX  
DR WPI; 1999-620430/53.  
XX  
PT New nucleic acid encoding sarcoma-associated gene products, useful for  
PT diagnosing, e.g. treating and preventing cancer  
XX  
PS Disclosure; Page 25; 93pp; English.  
XX  
CC The present invention describes sarcoma-associated gene products (I).  
CC Agents, specifically sarcoma associated nucleic acids (II) or their  
CC expression products that are tumour rejection antigens (TRA), that  
CC selectively increase formation of HLA (human leucocyte antigen)/(I)  
CC complexes are used for treating cancer, especially sarcoma and  
CC carcinoma, in humans and other animals. Compositions containing  
CC autologous cytolytic T cells (CTL), specific for the HLA/(I) complex,  
CC are similarly useful, also transformed cells that stimulate such CTL  
CC in vivo. (II) are also used; (I) as source of therapeutic antisen  
CC sequences that reduce expression of (II); (II) for recombinant  
CC production of (I); (III) particularly its fragments, as primers and  
CC probes in usual hybridisation and amplification assays, for diagnosis,  
CC prognosis and monitoring of tumours, or for measuring binding  
CC specificity of HLA molecules or CTL clones; (IV) to identify related  
CC sequences; and (V) for generating transgenic animals, e.g. for studying  
CC cancer and immune responses to it. (I) are used to raise specific  
CC antibodies (Ab) and therapeutically. Ab are used to diagnose tumours in  
CC immunosays, also for delivering drugs, toxins, imaging agents etc. to  
CC (I)-expressing cells. AAV49637 to AAV49670 represent exemplary tumour  
CC antigenic peptides given in the present invention.  
CC  
SQ Sequence 9 AA;  
  
Query Match 100.0%; Score 59; DB 20; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KTWGQYQYQV 9  
DB 1 ktwgqyqwv 9  
  
RESULT 8  
AAV53524  
ID AAV53524 standard; Protein; 9 AA.  
XX  
AC AAV53524;  
XX  
DT 18-JAN-2000 (first entry)  
XX  
DE Human melanoma Pmel17 (gp100) (aa 154-162) binds HLA-A2.  
XX  
KM Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;  
KM electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;  
KM human immunodeficiency virus; hepatitis B virus; papilloma virus;  
KM melanoma; malaria; parasite.  
XX  
OS Synthetic.  
XX  
PN FR2776926-A1.  
XX  
PD 08-OCT-1999.  
XX  
PF 07-APR-1998; 98FR-0004323.  
XX

PR 07-APR-1998; 98FR-0004323.  
 XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (CNRS ) CNRS CENT NAT RECH SCI.  
 PA (INSP ) INST PASTEUR LILLE.  
 PI Le Gal FA, Gullet JG, Gabery SH, Gras MH, Melnyk O, Tartar A;  
 DR WPI: 1999-583113/50.  
 XX  
 XX  
 PT New lipopeptide containing lipid regions and two epitopes, all  
 PT separated by peptide spacers that impart hydrophilicity, useful in  
 PT vaccines  
 XX  
 PS Disclosure; Page 24; 35pp; French.  
 XX  
 CC The invention relates to the generation of a lipopeptide comprising at  
 CC least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)  
 CC epitope and at least one lipid residue with (i) the epitopes and lipid  
 CC portion and (ii) the epitopes, being separated independently by peptide  
 CC spacers. These spacers comprise sequences of amino acids which carry an  
 CC overall electrical charge in neutral media to ensure that the  
 CC lipopeptide is hydrophilic. The peptides AAY3301-Y53549 represents  
 CC examples of peptide epitopes used to generate the lipopeptides. These are  
 CC used in therapeutic or prophylactic compositions and vaccines to induce  
 CC specific immune responses against human immunodeficiency, hepatitis B or  
 CC papilloma viruses; p53 of melanoma or the malaria parasite.  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 59; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. NO. 3.4e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTWGQYQV 9  
 Db 1 KTWGQYQV 9

RESULT 9  
 AAY47616  
 ID AAY47616 standard; Peptide; 9 AA.  
 XX  
 AC AAY47616;  
 XX  
 DT 01-DEC-1999 (first entry)  
 XX  
 DE Immunogenic peptide having a human leukocyte antigen binding motif #2227.  
 XX  
 KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;  
 KW immune response; T cell activation; major histocompatibility complex;  
 KW cytotoxic T lymphocyte; CTL; tumor rejection; viral infection; cancer;  
 KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;  
 KW vaccine; immunisation.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO945954-A1.  
 XX  
 PD 16-SEP-1999.  
 XX  
 PF 13-MAR-1998; 98WO-US05039.  
 XX  
 PR 13-MAR-1998; 98WO-US05039.  
 XX  
 PA (EPIM-) EPIMUNE INC.  
 XX  
 PI Sette A, Kubo RT, Sidney J, Cella E, Grey HM, Southwood S;  
 DR WPI: 1999-551214/46.  
 XX

PT New immunogenic peptides with HLA binding motif, useful in treatment  
 PT and diagnosis of cancers and viral diseases  
 PS Claim 1; Page 116; 150pp; English.  
 XX  
 CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides  
 CC having a human major histocompatibility complex (MHC) Class I (also  
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic  
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes  
 CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell  
 CC response against the antigen from which the peptide is derived.  
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are  
 CC normally induced by an antigen in the form of a peptide fragment bound  
 CC to a HLA molecule, rather than the intact foreign antigen itself, and  
 CC are particularly important in tumour rejection and in fighting viral  
 CC infections. The peptides are therefore useful therapeutically to treat  
 CC or prevent viral infections and cancers in mammals (especially humans)  
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.  
 CC They can be administered as vaccines to elicit an immune response in  
 CC individuals susceptible or otherwise at risk of viral infection or  
 CC cancer, or used to treat chronic or acute conditions. They are also  
 CC useful diagnostically, and can be used to induce a cytotoxic T cell  
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to  
 CC produce CTLs ex vivo for infusion back into a patient. The  
 CC polynucleotides encoding the immunogenic peptides are also useful  
 CC therapeutically and for immunisation as above.  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 59; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. NO. 3.4e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTWGQYQV 9  
 Db 1 KTWGQYQV 9

RESULT 10  
 AAY33172  
 ID AAY33172 standard; peptide; 9 AA.  
 XX  
 AC AAY33172;  
 XX  
 DT 17-NOV-1999 (first entry)  
 XX  
 DE Human gp100-Pmel117 peptide #1.  
 XX  
 KW Human; protein delivery; Yersinia sp.; effector gene; mutant; antigen;  
 KW immune response; cytotoxic T-lymphocyte; CTL; vaccination; treatment;  
 KW pathological disorder; gp100-Pmel117.  
 XX  
 OS Homo sapiens.  
 OS WO945098-A2.  
 XX  
 PN WO945098-A2.  
 XX  
 PD 10-SEP-1999.  
 XX  
 PF 03-MAR-1999; 99WO-IB00587.  
 XX  
 PR 06-MAR-1998; 98US-0036582.  
 XX  
 PA (VBRU/) VAN DER BRUGGEN P B.  
 PA (CORN/) CORNELIS G R.  
 PA (BOLA/) BOLAND A M.  
 PA (BOON/) BOON-FALEUR T R.  
 XX  
 PI Van Der Bruggen PB, Cornelis GR, Boland AM, Boon-Falleur TR;  
 DR WPI: 1999-540840/45.  
 XX  
 PT New mutant Yersinia strains useful for treating a pathological disorder

XX Example 1 : Page 71; 80pp; English.  
PS  
CC This invention describes a novel mutant Yersinia (Y1) strain, comprising  
CC of functional effector gene(s). The invention describes (1) a  
CC quintuple mutant Yersinia strain, having the designation Yersinia  
CC enterocolitica yopeHOMP or Yersinia pseudotuberculosis yopeHNOJ; (2) an  
CC expression vector (EV1) for delivering a heterologous protein into a  
CC eukaryotic cell, comprising in the 5'-3' direction: (3) a Yersinia or  
CC mutant Yersinia strain for delivering a heterologous protein into a  
CC eukaryotic cell, comprising contacting the cell with a Y1 transformed  
CC with the above vector (Y1-EV1); (4) a method for delivering a  
CC heterologous protein into a eukaryotic cell, comprising contacting the  
CC cell with a Y1 transformed with the above vector (Y1-EV1); (5) a method  
CC for inducing an immune response specific for a heterologous protein; (6)  
CC a method for inducing a cytotoxic T-lymphocyte (CTL) response specific  
CC for a heterologous protein; (7) a method for determining the efficacy of  
CC an antigen vaccination regimen in a subject. Y1 is used to treat a  
CC pathological disorder, by providing recombinant Yersinia for the safe  
CC delivery of proteins into eukaryotic cells. AAY3147-Y33178 are  
CC human-derived peptides used to illustrate the method of the invention.  
SO Sequence 9 AA:

Query Match 100.0%; Score 59; DB 20; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
Db 1 ktwgqyqv 9

## RESULT 11

AAY40211  
ID AAY40211 standard; Peptide; 9 AA.

XX AAY40211;

DT 19-NOV-1999 (first entry)

DE Amino acid sequence of a human melanoma epitope.

KW Cytotoxic T cell; T lymphocyte; CD8+ epitope; T helper cell;

KW CD4+ epitope; B epitope; lipopeptide; interferon gamma; adjuvant;

KW vaccine; tumor; infection; immune response; cytokine profile;

KW acquired immune deficiency syndrome; papilloma; cancer; hepatitis;

XX autoimmune disease.

OS Homo sapiens.

PN FR2774687-A1.

PD 13-AUG-1999.

PF 06-FEB-1998; 98FR-0001439.

PR 06-FEB-1998; 98FR-0001439.

PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

PI (INSP ) INST PASTEUR LILLE.

DR Thiam K, Guillet JG, Ver Waerde C, Auriault C, Gras MH, Loinig E;

XX WPI; 1999-510734/43.  
XX  
PT New lipopeptide comprising C-terminal interferon-gamma fragment with  
PT attached lipophilic groups, used as interferon mimic, e.g. for treating  
PS cancer or virus infection

PS Disclosure; Page 35; 53pp; French.

XX AAY40123-Y40379 represent epitopes that are able to activate cytotoxic  
CC T lymphocytes (CD8+ epitopes), T helper cells (CD4+ epitopes), or  
CC B epitopes recognized by corresponding antibodies. The epitopes may be  
CC used in the composition of the invention. The specification describes a  
CC lipopeptide that has a peptide part derived from mammalian interferon  
CC gamma (IFN $\gamma$ ) and one or more lipophilic parts comprising a linear or  
CC branched, (un)saturated 4-20C hydrocarbon chain or a steroid. The  
CC lipopeptide mimics the activity of IFN $\gamma$ . Compositions comprising the  
CC lipopeptide are used to treat or prevent any condition that responds  
CC to IFN $\gamma$ , and as adjuvant for vaccines (particularly those directed  
CC against tumors, viral or parasitic infections), to stimulate or  
CC (re)orient the immune response between types 1 and 2 cytokine profiles.  
CC Particular applications are treatment of infections (particularly  
CC viral, e.g. acquired immune deficiency syndrome, papilloma (cancer) and  
CC hepatitis) but also bacterial, fungal, parasitic or helminth; cancers  
CC (particularly of kidney, cutaneous T cells or ovary, chronic  
CC myelogenous leukemia or mesothelioma), allergy; and autoimmune  
CC diseases.  
SO Sequence 9 AA:

Query Match 100.0%; Score 59; DB 20; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
Db 1 ktwgqyqv 9

## RESULT 12

AAY26867  
ID AAY26867 standard; peptide; 9 AA.

XX AAY26867;

DT 14-SEP-1999 (first entry)

DE Melanoma-derived lipopeptide epitope #8 for mixed micelles.

KW Micelle; microaggregate; induction; immune response; lipopeptide; CTL;

KW cytotoxic T-lymphocyte; epitope; lipid; helper T-lymphocyte; HTV; HBV;

KW tetanus; toxin; vaccine; HIV; hepatitis B virus; papilloma virus; p53;

XX melanoma; Plasmodium falciparum; malaria.

OS Synthetic.

OS Homo sapiens.

PN FR2771640-A1.

PD 04-JUN-1999.

PF 03-DEC-1997; 97FR-0015246.

PR 03-DEC-1997; 97FR-0015246.

PA (CNRS ) CNRS CENT NAT RECH SCI.

PI (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

PI (INSP ) INST PASTEUR LILLE.

PI Bossus M, Bourgaud VI, Gras-Masse H, Guillet JG, Lippens G;

XX Tatar A, Wieruszski JM;

XX WPI; 1999-349509/30.  
XX  
PT Immunogenic lipopeptide micelles - comprising lipopeptides  
PT containing cytotoxic and helper T-lymphocyte epitopes  
PS  
PS Disclosure; Page 37; 60pp; French.

PS The invention relates to the generation of mixed micelles or

CC microaggregates for inducing an immune response comprise: (a) a first  
CC lipopeptide comprising at least one CTL (cytotoxic T-lymphocyte) epitope  
CC and at least one lipid unit; and (b) a second lipopeptide comprising at  
CC least one HTL (helper T-lymphocyte) epitope and at least one lipid unit  
CC different from that of the first lipopeptide. This peptide represents  
CC an example of a lipopeptide epitope used in the invention and is derived  
CC from a human melanoma protein. The immunogenic lipopeptide micelles  
CC are used in vaccines, especially against HIV, hepatitis B virus (HBV),  
CC papilloma viruses, p53, melanoma or Plasmodium falciparum malaria.

XX Sequence 9 AA:

Query Match 100.0%; Score 59; DB 20; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMQV 9  
Db 1 KTWGQYMQV 9

RESULT 13  
AAY01753 standard; Peptide: 9 AA.

XX AAY01753:

XX 25-JUN-1999 (first entry)

XX Exemplary antigenic peptide derived from gp100(Pmel117).

XX MAGE-3: tumour associated gene; human leucocyte antigen Class II;

XX autologous CD4+ cell; MAGE-3 related disease; cancer; melanoma;

XX osteosarcoma; leukemia; carcinoma.

XX Homo sapiens.

XX WO914326-A1.

XX 25-MAR-1999.

XX 04-SEP-1998; 98WO-US18601.

XX 12-SEP-1997; 97US-0928615.

XX (LUDM-) LUDWIG INST CANCER RES.

XX (UYVR-) UNIV VIRJE BRUSSEL.

XX Boon-Falleur T, Chaux P, Cortals J, Heirman C;

XX Luiten R, Stroobant V, Thielemans K, Van Der Bruggen P;

XX WPI: 1999-244031/20.

XX Isolated peptides that bind to human leucocyte antigen class II

XX molecules

XX Disclosure: Page 29; 88pp; English.

XX The present sequence represents an exemplary tumour associated peptide

XX antigen. The specification describes a MAGE-3 tumour associated gene.

XX Peptides (AA01721-25) that bind human leucocyte antigen (HLA) Class II

XX molecules can be derived from the MAGE-3 protein. These peptides and

XX autologous CD4+ cells that bind to a complex of MAGE-3 peptide

XX and HLA Class II, are used to treat MAGE-3 related diseases,

XX particularly cancers (e.g. melanoma, osteosarcoma, leukemia and

XX various forms of carcinoma). The peptides are also used to produce

XX specific antibodies. Detection of the peptides, e.g. in binding

XX assays, particularly with antibodies, is used for diagnosis of such

XX diseases.

XX Sequence 9 AA:

Query Match 100.0%; Score 59; DB 20; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMQV 9  
Db 1 KTWGQYMQV 9

RESULT 14  
AAY00715 standard; peptide: 9 AA.

XX AAY00715:

XX 12-MAY-1999 (first entry)

XX Tumour antigen booster peptide gp100Pmel117 HLA-A2 #1.

XX Tumour antigen; booster peptide; immune response modulation; allergy;

XX immune response enhancer; tumour cell; tumour rejection antigen;

XX leukocyte antigen-presenting molecule; autoimmune disease;

XX allograft rejection.

XX Homo sapiens.

XX WO9858956-A2.

XX 30-DEC-1998.

XX 19-JUN-1998; 98WO-US12894.

XX 23-JUN-1997; 97US-0880979.

XX (LUDM-) LUDWIG INST CANCER RES.

XX Boon-Falleur T, Uytendhoeve C, Warnier G;

XX WPI: 1999-105612/09.

XX Immunization methods using viruses expressing antigen for priming

XX and booster immunizations - useful for modulating immune responses

XX against antigen, e.g. enhancing immune response against tumour cells

XX expressing tumour rejection antigens

XX Disclosure: Page 10; 33pp; English.

XX This sequence represents a tumour antigen booster peptide that can be

XX used in the method of the invention. The method is for modulating an

XX immune response in a mammal against an antigen, and comprises:

XX (A) inducing an immune response by: (i) administering a virus containing

XX a nucleic acid molecule encoding the antigen or its precursor to generate

XX an immune response; and (ii) administering at least one booster dose

XX comprising a peptide including the antigen, in an adjuvant, in a combined

XX amount effective to enhance the initial immune response; or

XX (B) reducing an immune response as defined for (A) but using a

XX non-adjuvant with the peptide which includes the antigen, in an amount

XX effective to reduce the initial immune response. Method (A) is used to

XX enhance the immune response against tumour cells expressing tumour

XX rejection antigens, and against pathogens in subjects having human

XX leukocyte antigen-presenting molecules. Method (B) is used to reduce the

XX immune response in allergy, autoimmune disease, and allograft rejection.

XX Method (A) provides an immunisation method which, unlike prior art, is

XX not limited by the host immune response against viral vectors.

XX Sequence 9 AA:

Query Match 100.0%; Score 59; DB 20; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
 ID |||||||  
 DB 1 ktwgqyqv 9

RESULT 15  
 AAY10449

ID AAY10449 standard; Peptide; 9 AA.

AC AAY10449;

DT 12-MAY-1999 (first entry)

DE HLA Class I motif peptide SEQ ID NO:379.

KW Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;  
 immunisation; tumour; infectious disease; immunotherapy; cancer;  
 malignant melanoma; viral disease; hepatitis; AIDS.

OS Synthetic.

OS Homo sapiens.

PN MO9902183-A2.

PD 21-JAN-1999.

PF 10-JUL-1998; 98WO-US14289.

PR 10-DEC-1997; 97US-0988320.

PR 10-JUL-1997; 97CA-2209815.

PA (CTL-) CTL IMMUNOTHERAPIES CORP.

PI Kuendig TM, Simard JTL;

DR WPI: 1999-120514/10.

PT Inducing a cytotoxic T lymphocyte response - by maintaining a level  
 of antigen in the lymphatic system of a mammal so as to provide a  
 sustained CTL response, used to treat, e.g. AIDS

PS Disclosure; Page 40; 199pp; English.

CC The present invention describes a method of inducing and/or sustaining  
 an immunological cytotoxic T lymphocyte (CTL) response in a mammal. The  
 method comprises: (a) delivering an antigen to the mammal at a level to  
 induce an immunological CTL response in the mammal; and (b) maintaining  
 the level of the antigen in the mammal's lymphatic system to maintain  
 the immunologic CTL response. The method can be used for the delivery of  
 e.g. a differentiation antigen, a tumour-specific multilineage antigen,  
 an embryonic antigen, an oncogene antigen, a mutated tumour-suppressor  
 gene antigen, or a viral antigen. They can be used for the treatment of  
 disease such as cancer, e.g. malignant melanoma or infectious disease,  
 e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery  
 to the lymphatic system provides for potent CTL stimulation that takes  
 place in the milieu of the lymphoid organ, and it sustains stimulation  
 that is necessary to keep CTL active, cytotoxic and recirculating  
 through the body. AAY10071 to AAY10639 represent examples of peptide  
 antigens given in the present invention.

CC Sequence 9 AA;

Query Match 100.0%; Score 59; DB 20; Length 9;

Best Local Similarity 100.0%; Pred. NO. 3.4e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
 ID |||||||  
 DB 1 ktwgqyqv 9

RESULT 16

AAB33662  
 ID AAB33662 standard; Peptide; 9 AA.

AC AAB33662;

DT 26-JAN-2001 (first entry)

DE MHC class I associated immunogenic peptide SEQ ID 61.

KW Microparticle; nucleic acid delivery; immunogenic peptide; MHC I; MHC II;  
 major histocompatibility complex; vaginal tissue; mucosal tissue..

OS Unidentified.

PN WO200053161-A2.

PD 14-SEP-2000.

PF 10-MAR-2000; 2000WO-US06578.

PR 11-MAR-1999; 99US-0266463.

PR 27-MAY-1999; 99US-0321346.

PA (ZYCO-) ZYCO INC.

PI Lunsford IB, Putnam D, Hedley ML;

DR WPI: 2000-638130/61.

PT Microparticles useful for administering a nucleic acid into the mucosal  
 tissue preferably vaginal tissue of an animal, comprises a polymeric  
 matrix, a lipid and a nucleic acid molecule

PS Disclosure; Page 14; 96pp; English.

CC The present invention relates to microparticles which are less than 20  
 microns in diameter, which comprise a polymeric matrix, a lipid and a  
 nucleic acid molecule. The microparticle is specifically not  
 encapsulated in a liposome and does not comprise a cell. The nucleotide  
 sequence encodes an expression product that binds to major  
 histocompatibility complex (MHC) type I or II molecules. Peptides  
 CC AAB33602-B33647 represent MHC class II associated immunogenic peptides,  
 and AAB33648-B33710 represent MHC class I associated immunogenic  
 peptides. The peptides are examples of the expression products of the  
 CC nucleotide sequences which can be included in the microparticles of the  
 CC invention. Sequences AAB33711-B33716 represent alternative expression  
 CC products and nuclear localisation signals also used in the invention. The  
 CC microparticles are useful for administering a nucleic acid into the  
 CC mucosal tissue preferably vaginal tissue of an animal.

CC Sequence 9 AA;

Query Match 100.0%; Score 59; DB 21; Length 9;

Best Local Similarity 100.0%; Pred. NO. 3.4e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
 ID |||||||  
 DB 1 ktwgqyqv 9

RESULT 17  
 AAB23679

ID AAB23679 standard; Peptide; 9 AA.

AC AAB23679;

DT 05-JAN-2001 (first entry)

DE Cytotoxic T lymphocyte (CTL) epitope SEQ ID NO:31.  
 KW ATPase; Hsp70; heat shock protein; cytotoxic T lymphocyte; CTL;

KW Immune response; infectious disease; malaria; cytotoxic T cell;  
KW cytostatic; immunostimulant; cellular immune response inducer;  
KW protozoacide; leukaemia; cancer.  
XX  
OS Homo sapiens.  
PN WO200049041-A1.  
XX  
PD 24-AUG-2000.  
XX  
PF 18-FEB-2000; 2000WO-JP00941.  
XX  
PR 19-FEB-1999; 99JP-0041535.  
XX  
PA (SUME) SUMITOMO ELECTRIC IND CO.  
XX  
PI Shinbara N, Udono H, Yui K;  
XX  
DR WPI; 2000-543748/49.  
XX  
PT Fused protein capable of inducing cellular immune response, useful as  
PT active ingredient for drug compositions in preventing and/or treating  
PT infectious diseases such as malaria or cancer  
XX  
PS Claim 7; Page 58; 72pp; Japanese.  
XX  
CC The present invention describes a fused protein (I) prepared from a  
CC peptide containing a CTL (cytotoxic T lymphocyte) epitope recognised by  
CC cytotoxic T cells and a protein containing the ATPase domain of a heat  
CC shock protein. Also described are: (1) a drug composition containing (I)  
CC as active ingredient; (2) a DNA encoding (I); (3) an expression vector  
CC containing the DNA of (2); and (4) a transformant which can retain the  
CC expression vector of (3). (I) has cytostatic, immunostimulant and  
CC protozoacide activities, and can be used as a cellular immune response  
CC inducer. The protein is useful as an active ingredient for drug  
CC compositions in preventing and/or treating infectious diseases such as  
CC malaria or cancer e.g. to provide systemic immunity against leukaemia.  
CC The present sequence represents a specifically claimed CTL epitope  
CC for use in a fused protein of the present invention.  
XX  
SQ Sequence 9 AA:  
  
Query Match 100.0%; Score 59; DB 21; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KTWGQYMOV 9  
Db 1 ktwgqygqv 9  
|||||  
1 ktwgqygqv 9  
  
RESULT 18  
AAB08694  
ID AAB08694 standard; Peptide; 9 AA.  
XX  
AC AAB08694;  
XX  
DT 02-JAN-2001 (first entry)  
XX  
DE Antigenic peptide from tumour rejection antigen gp100(Pmel117).  
XX  
KW Epha3; HLA class II-binding peptide; human leukocyte antigen; antigen;  
KW CD4+ T lymphocyte; tumour associated gene; vaccine.  
XX  
OS Homo sapiens.  
XX  
PN WO200050589-A1.  
XX  
PD 31-AUG-2000.  
XX  
PF 18-FEB-2000; 2000WO-US04326.  
XX

PR 22-FEB-1999; 99US-0121170.  
PR 08-OCT-1999; 99US-0158566.  
XX  
XX (LUDW-) LUDWIG INST CANCER RES.  
XX  
PI Chari R, Coulle P, Boon-Falleur T;  
XX  
DR WPI; 2000-572089/53.  
XX  
XX Novel tyrosine kinase receptor, Epha3 human leukocyte antigen (HLA)  
PT class II binding peptide and nucleic acid encoding the receptor, useful  
PT for diagnosing and treating conditions characterized by expression of  
PT Epha3 gene  
XX  
XX Disclosure; Page 35; 107pp; English.  
XX  
XX AAB08668-B08704 represent antigenic peptides characteristic of tumours.  
CC The peptides may be combined in vaccines with a human Epha3 HLA (human  
CC leukocyte antigen) class II-binding peptide. Epha3 antigens, when  
CC presented by an antigen presenting cell having a HLA class II molecule,  
CC effectively induce activation and proliferation of CD4+ T lymphocytes.  
CC Epha3 is a tumour associated gene. Epha3 HLA binding peptides are used  
CC for selectively enriching a population of T lymphocytes. The peptides  
CC are also used for diagnosing a disorder characterized by Epha3 or Epha3  
CC HLA binding peptide expression. The peptides are also used to treat a  
CC disorder characterized by Epha3 expression. The Epha3 binding peptides  
CC are useful in producing vaccines and antibody.  
XX  
SQ Sequence 9 AA:  
  
Query Match 100.0%; Score 59; DB 21; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KTWGQYMOV 9  
Db 1 ktwgqygqv 9  
|||||  
1 ktwgqygqv 9  
  
RESULT 19  
AAY71520  
ID AAY71520 standard; peptide; 9 AA.  
XX  
AC AAY71520;  
XX  
DT 12-OCT-2000 (first entry)  
XX  
DE Human gp100 Pmel117 nonapeptide-1.  
XX  
KW gp100 Pmel117; human; Tumour Rejection Antigen; TRA; tumour; cancer;  
KW HLA; Human Leucocyte Antigen; MHC; Major Histocompatibility Complex; CTL;  
KW cytolytic T-lymphocyte; immune response stimulator; prophylaxis; therapy;  
KW diagnosis; TNF; tumour necrosis factor; vaccine; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO200032769-A2.  
XX  
PD 08-JUN-2000.  
XX  
PF 26-NOV-1999; 99WO-IB02018.  
XX  
PR 27-NOV-1998; 98GB-0026143.  
XX  
PA (LUDW-) LUDWIG INST CANCER RES.  
XX  
PI Huang L, Van Pel A, Brasseur F, De Plaen E, Boon T;  
XX  
DR WPI; 2000-412317/35.  
XX  
PT Novel polypeptides expressed in tumor cells useful for treating cancers  
PT have an ability to complex with a major histocompatibility complex

PT molecule and comprises a specific unbroken amino acid sequence  
 XX  
 PS Disclosure: Page 20; 80pp; English.  
 XX  
 CC The patent discloses MAGE-A10 and MAGE-A8 polypeptide, nonapeptide and  
 CC decapeptide sequences, that function as tumour rejection antigens  
 CC (TRAs). These peptides are capable of forming a complex with major  
 CC histocompatibility complex (MHC) molecule type HLA-A2.1 (Human Leucocyte  
 CC Antigen), that are recognised by T-lymphocytes and elicit an immune  
 CC response from cytolytic T-lymphocytes (CTL). They function as an immune  
 CC response stimulator. Tumour rejection antigens are useful in prophylaxis,  
 CC therapy and diagnosis of tumours and are effective in controlling or  
 CC preventing tumour growth. The present sequence is the human gp100 Pmel117  
 CC nonapeptide-1, that corresponds to residues 154-162 of the tumour  
 CC associated gene, gp100 Pmel117 encoding protein. It can be administered  
 CC to induce or enhance an immune response and is presented by HLA-A2  
 CC complex. This peptide can serve as a tumour rejection antigen (TRA) and  
 CC in combination with adjuvants, can produce vaccines useful for treating  
 CC a variety of tumours.  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 59; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. NO. 3.4e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
 ||||||||  
 DB 1 ktwgqyqv 9

## RESULT 20

AAB02622  
 ID AAB02622 standard; Peptide; 9 AA.

AC AAB02622;

DT 18-AUG-2000 (first entry)

XX Tumour associated peptide antigen from gp100 #2.

DE MAGE-A3; HLA class II; human leukocyte antigen; antibody; vaccine;  
 XX cancer; human; tumour; tumour associated gene product.

XX Homo sapiens.

XX WO200020581-A1.

XX 13-APR-2000.

XX 15-SEP-1999; 99WO-US21230.

XX 05-OCT-1998; 98US-0166448.

XX (LUDM-) LUDMIG INST CANCER RES.

XX (UYVR-) UNIV VRIJE BRUSSEL.

XX Chaux P, Stroobant V, Boon-Falleur T, Van Der Bruggen P;  
 PI Schultz ES, Van Snick J, Lethe B, Thielemans K, Cortals J;  
 PI Heitman C;

XX WPI: 2000-317713/27.

PT New MAGE-A3 class II binding peptides, useful to diagnose and treat  
 PT tumours, are fragments of MAGE-A3 which bind to and are presented to T  
 PT lymphocytes by human leukocyte antigen class II molecules  
 XX  
 PS Disclosure: Page 33; 119pp; English.

XX The present invention relates to MAGE-A3 (tumour associated gene  
 CC product) human leukocyte antigen (HLA) class II-binding peptides (see  
 CC AAB02566-B02565, and AAB02633-B02637). These peptides are presented to T

CC cells in the context of HLA class II molecules. The peptides stimulate  
 CC the activity and proliferation of CD4+ T lymphocytes. The invention also  
 CC includes nucleotide sequences encoding MAGE-3A peptides (see AAA37928  
 CC and AAA37938-A37940). The peptides and nucleotide sequences can be used  
 CC to create antibodies against the MAGE-A3 peptides, the antibodies, the  
 CC peptides and nucleotide sequences can be used to create a vaccine. The  
 CC peptides are used to diagnose or treat a disorder characterized by  
 CC expression of MAGE-3, particularly cancer. The methods can also be used  
 CC in the diagnosis of disorders associated with MAGE-3 expression. Included  
 CC in the invention are other human tumour antigens (see AAB02596-B02637),  
 CC and PCR primers used in the course of the invention (see AAA37929-A37937  
 CC and AAA37941-A37942).

Query Match 100.0%; Score 59; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. NO. 3.4e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
 ||||||||  
 DB 1 ktwgqyqv 9

## RESULT 21

AA90803  
 ID AA90803 standard; peptide; 9 AA.

AC AA90803;

DT 25-AUG-2000 (first entry)

XX Human leukocyte antigen A2 gp100-Pmel117 peptide SEQ ID NO:32.

DE Human leukocyte antigen; HLA-B35; binding; recognition; lysis;  
 XX cytolytic T cell; tyrosinase; immune response; diagnosis;  
 XX identification; human.

XX Homo sapiens.

XX WO200021551-A1.

XX 20-APR-2000.

XX 04-OCT-1999; 99WO-US23038.

XX 09-OCT-1998; 98US-0169717.

XX (LUDM-) LUDMIG INST CANCER RES.

XX Ooms A, De Giovanni G, Morel S, Van Den Eynde B, Boon-Falleur T;  
 PI WPI: 2000-317842/27.

PT Isolated peptides, sometimes derived from tyrosinase, which bind to  
 PT HLA-B35 leading to recognition and lysis of the resulting complexes by  
 PT cytolytic T cells  
 XX  
 PS Example 3; Page 9; 20pp; English.

XX The present invention describes isolated peptides which bind to human  
 CC leukocyte antigen (HLA)-B35 molecules leading to recognition and lysis  
 CC of the resulting complexes by cytolytic T cells. The isolated peptides  
 CC are sometimes derived from tyrosinase. Compositions comprising the  
 CC peptides of the invention can be used to generate immune responses,  
 CC preferably in humans, but also in non-human animals to generate immune  
 CC components which can then be used to treat humans or diagnostically.  
 CC Therapeutically, the peptides are useful in generation of cytolytic T  
 CC cells either in vitro or in vivo which specifically lyse pathogenic  
 CC cells. The peptides can also be used to identify HLA-B35 positive  
 CC cells, or to remove HLA-B35 positive cells from mixtures containing  
 CC such cells. Nucleic acid molecules encoding the peptides can be used



CC Inter alia as probes to identify cells which are expressing tyrosinase.  
 CC The present sequence represents an HLA binding peptide used in the  
 CC exemplification of the present invention.

XX  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 59; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
 1 ktwgqywgqv 9

RESULT 22

AAV92299  
 ID AAV92299 standard; peptide: 9 AA.

XX  
 AC AAV92299;

XX  
 DT 10-AUG-2000 (first entry)

XX  
 DE gp100-Pmel117 antigenic peptide epitope (residues 154-162).

XX  
 KW gp100-Pmel117; antigen: epitope; cytotoxic T lymphocyte; CTL; complex;

XX  
 KW human leukocyte antigen; HLA.

XX  
 OS Homo sapiens.

XX  
 PN WO200020445-A2.

XX  
 PD 13-APR-2000.

XX  
 PF 15-SEP-1999; 99WO-IB01664.

XX  
 PR 02-OCT-1998; 98US-0165863.

XX  
 PR 09-APR-1999; 99US-0289350.

XX  
 PA (CHAU/) CHAUX P.

XX  
 PA (LUIT/) LUITEN R.

XX  
 PA (DEMO/) DEMOTE N.

XX  
 PA (DUFF/) DUFFOUR M.

XX  
 PA (LURQ/) LURQUIN C.

XX  
 PA (TRAY/) TRAVERSARI C.

XX  
 PA (STRO/) STROOBANT V.

XX  
 PA (CORN/) CORNELIS G. R.

XX  
 PA (BOON/) BOON-FALLEUR T.

XX  
 PA (YBRU/) VAN DER BRUGGEN P.

XX  
 PI CHAUX P, Luiten R, Demotte N, Duffour M, Lurquin C, Traversari C;

XX  
 PI Stroobant V, Cornelis GR, Boon-Falleur T, Van Der Bruggen P;

XX  
 PI Schults E, Warnier G;

XX  
 DR WPI; 2000-303739/26.

XX  
 PT Isolation of cytotoxic T-lymphocytes clones by successive steps of

XX  
 PT stimulation and testing of lymphocytes with antigen presenting cells

XX  
 PT which present antigens derived from different expression systems

XX  
 PS Disclosure; Page 22; 99pp; English.

XX  
 CC A novel method of isolation of cytotoxic T-lymphocytes (CTL) clones

XX  
 CC comprises successive steps of stimulation and testing of lymphocytes

XX  
 CC with antigen presenting cells (APCs) which present antigens derived

XX  
 CC from different expression systems. The CTL clones isolated recognize

XX  
 CC specific antigenic peptides of proteins, preferably of the MAGE family.

XX  
 CC The APC is autologous and each expression system is different from at

XX  
 CC least one of the other expression systems, therefore isolating a

XX  
 CC cytotoxic T cell clone specific for the protein. The method can also be

XX  
 CC used to identify an antigenic peptide epitope. Isolated CTL clones

XX  
 CC specific for a peptide/human leukocyte antigen (HLA) complex are claimed.

CC The CTL cells specific for the complexes, peptides or cells which present  
 CC the complexes on the cell surface are useful for treating pathological  
 CC conditions characterized by abnormal expression of the complexes.

XX  
 XX  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 59; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
 1 ktwgqywgqv 9

RESULT 23

AAV84296  
 ID AAV84296 standard; peptide: 9 AA.

XX  
 AC AAV84296;

XX  
 DT 12-JUL-2000 (first entry)

XX  
 DE Tumour associated antigen derived from gp100-Pmel117.

XX  
 KW tumour rejection antigen; macrophage colony stimulating gene;

XX  
 KW macrophage-colony stimulating factor; antigen presenting cell;

XX  
 KW human leukocyte antigen; CD8+ cytotoxic T lymphocyte.

XX  
 OS Homo sapiens.

XX  
 PN WO200013699-A1.

XX  
 PD 16-MAR-2000.

XX  
 PF 03-SEP-1999; 99WO-US20344.

XX  
 PR 04-SEP-1998; 98US-0099077.

XX  
 PA (LUDW-) LUDWIG INST CANCER RES.

XX  
 PA Probst-Keeper M, Van Den Eynde B, Boon-Falleur T;

XX  
 PI WPI; 2000-256859/22.

XX  
 DR Isolated polypeptide used to treat subjects having a disorder

XX  
 PT characterized by expression of alternative open reading frame

XX  
 PT macrophage-colony stimulating factor comprises 25 amino acid residue

XX  
 PT sequence -

XX  
 PS Disclosure; Page 21; 74pp; English.

XX  
 CC AAV84270-Y84303 represent peptides which are tumour associated antigens.

XX  
 CC They can be administered in conjunction with the tumour rejection antigen

XX  
 CC precursor of the invention to induce anti-tumour responses. The tumour

XX  
 CC rejection antigen precursor of the invention is encoded by an

XX  
 CC alternative open reading frame (ORF) of human macrophage colony

XX  
 CC stimulating gene. Peptides derived from the alternative ORF of

XX  
 CC macrophage-colony stimulating factor, when presented by an antigen

XX  
 CC presenting cell having a human leukocyte antigen (HLA) class I molecule,

XX  
 CC effectively induce the activation and proliferation of CD8+ cytotoxic T

XX  
 CC lymphocytes. Polypeptide and nucleic acids derived from the alternate ORF

XX  
 CC of macrophage-colony stimulating factor are useful for enriching

XX  
 CC selectively a population of T lymphocytes with CD8+ T lymphocytes. They

XX  
 CC are also useful for diagnosing a disorder characterized by expression of

XX  
 CC the polypeptide, and for identifying functional variants and mimetics.

XX  
 SQ Sequence 9 AA;

XX  
 Query Match 100.0%; Score 59; DB 21; Length 9;

XX  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTMGOYMOV 9  
| | | | | | | | | |  
Db 1 KTMGYQYQV 9

RESULT 24  
AAV82979  
ID AAV82979 standard; Peptide; 9 AA.

AAV82979;

19-JUN-2000 (first entry)

gp100(Pmel117) tumour associated antigen.

MM Tumour; tumour associated antigen; retrovirus; antisense;  
KW treatment; probe; primer; HLA; cytotoxic T-lymphocyte; cancer;  
KW testis; antibody; CTL; helper T-lymphocyte; MAGE; GAGE; RAGE;  
KW Gnt-V; MDM; CDK4; beta catenin; tyrosinase; Melan-A; gp100; PRAME.

OS Homo sapiens.

PN WO200006598-A1.

10-FEB-2000.

15-JUL-1999; 99WO-US16236.

29-JUL-1998; 98US-0124398.

(LUDW-) LUDWIG INST CANCER RES.

Coulie P, Boon-falleur T;

WPI: 2000-205453/18.

Novel nucleic acids encoding melanoma associated gene products and their fragments and variants, useful for treating endogenous retrovirus mediated tumors, especially melanomas

Disclosure; Page 26; 77pp; English.

CC Tumors associated disorders (e.g. endogenous retrovirus mediated  
CC tumors, especially melanomas) can be treated or ameliorated by  
CC administering antisense nucleic acid to reduce the expression of  
CC tumor associated genes such as HERV-AVL3-B. Progression of  
CC a disorder characterized by the expression of the HERV-AVL3-B  
CC endogenous retrovirus tumor rejection antigen (ERTRA) can be  
CC diagnosed or monitored by contacting a non-testis biological  
CC sample with an agent that binds to the complex and determining  
CC the interaction. A disorder can also be treated by administering  
CC an agent that enriches the presence of HLA and HERV-AVL3-B ERTA  
CC or by administering autologous cytotoxic T-cells sufficient to  
CC ameliorate the disorder. Fragments of the HERV-AVL3-B coding sequence  
CC are useful as probes or amplification primers for determining the  
CC expression of HERV-AVL3-B genes, to express tumor associated  
CC polypeptides in vivo and in vitro and to prepare fragments of such  
CC polypeptides to synthesize antibodies. Antigenic peptides of  
CC HERV-AVL3-B can be useful for generating antibodies either alone or  
CC as fusion proteins, as components of immunoassay and for determining  
CC the binding specificity of HLA molecules and/or cytotoxic T  
CC lymphocyte (CTL) for HERV-AVL3-B proteins. Peptides derived from  
CC the HERV-AVL3-B coding sequence and which are presented by MHC  
CC molecules and recognised by CTL or helper T-lymphocytes can be  
CC combined with peptides from other tumour rejection antigens by  
CC preparation of hybrid nucleic acids or polypeptides to produce  
CC polytopes. This exemplary tumour associated peptide antigen  
CC corresponds to amino acids 154-162 of the gp100(Pmel117) polypeptide.  
CC See also AAV82953-Y82986.

Sequence 9 AA;

Query Match 100.0%; Score 59; DB 21; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTMGOYMOV 9  
| | | | | | | | | |  
Db 1 KTMGYQYQV 9

RESULT 25  
AAV56614

ID AAV56614 standard; peptide; 9 AA.

AAV56614;

21-FEB-2000 (first entry)

gp100-Pmel-117 gene MHC molecule HLA-A2 peptide SEQ ID NO:30.

MM HLA-A\*0201; human leukocyte antigen; cytolytic T cell; CTL; tumour;  
KW Melan-A; peripheral blood lymphocyte; FBL; immune complex; melanoma;  
KW MHC molecule; beta2-microglobulin; cytotoxic T lymphocyte; vaccine;  
KW immune response; cancer; tyrosinase; tumour rejection antigen;  
KW major histocompatibility complex.

OS Synthetic.

PN WO9950637-A2.

07-OCT-1999.

25-MAR-1999; 99WO-US06615.

27-MAR-1998; 98US-0049850.

(LUDW-) LUDWIG INST CANCER RES.

(UYOX-) UNIV OXFORD.

Romero P, Dunbar R, Valmori D, Ogg G, Cerrotini J, Cerundolo V;

WPI: 2000-052636/04.

New isolated complex of binding partners and immune complexes containing major histocompatibility molecules and peptide, used to isolate and detect cytotoxic T cells, particularly directed against cancer

Example 50; Page 64; 91pp; English.

CC The present invention describes an isolated complex (A) comprising: (i)  
CC first and second binding partners (BP1, BP2); and (ii) several immune  
CC complexes (IC) containing a major histocompatibility complex (MHC)  
CC molecule (I), a beta2-microglobulin molecule (b2m) and a peptide (II)  
CC that binds specifically to (I). (A) are used for analysis of cytolytic  
CC T cells (CTL) for characterisation of an immune response to tumours or  
CC for monitoring vaccine trials. Particularly they are used to isolate or  
CC detect particular CTL (especially those in tumour-infiltrated lymph  
CC nodes), including visualisation of antigen-specific CTL and determination  
CC if the cells have been activated by in vivo exposure to antigen. Isolated  
CC precursor cells may be expanded in vitro to produce cells with high  
CC tumoricidal activity, for therapeutic or diagnostic use. A method from  
CC the present invention allows: (i) preselection of T cell clones for use  
CC in immunotherapy according to their homing molecules; and (ii) improves  
CC the lytic activity of T cells populations by inhibition of natural  
CC killer cell receptors. The present sequence represents a peptide  
CC used in the exemplification of the present invention.

Sequence 9 AA;

Query Match 100.0%; Score 59; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYQV 9  
 |||||||  
 1 ktwgqyqgv 9

RESULT 26  
 AAB31354

ID AAB31354 standard; peptide; 9 AA.

XX AAB31354;

XX 20-APR-2001 (first entry)

XX Exemplary antigen characteristic of tumours, derived from gp100.

XX MAGE-A1: HLA: human leukocyte antigen; CD4+ T lymphocyte; cancer;

KM MAGE-A1 HLA class II-binding protein; vaccine.

XX Homo sapiens.

XX WO200078806-A1.

XX 28-DEC-2000.

XX 14-JUN-2000; 2000WO-US16287.

XX 18-JUN-1999; 99US-0336091.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Van Snick J, Lethe B, Chaux P, Boon-Falleur T, Van Der Bruggen P;

XX WPI; 2001-102698/11.

XX Novel MAGE-A1 human leukocyte antigen class II peptides which bind to  
 PT and are presented to the class II molecules, useful for inducing immune  
 PT response and treating cancers characterized by expression of MAGE-A1 -

XX Disclosure; Page 33; 78pp; English.

XX AAB31302-59 represent exemplary antigens which are characteristic

XX of tumours. They can be used to enhance the immune response of vaccines

XX comprising peptides derived from human MAGE-A1 HLA (human leukocyte

XX antigen) class II-binding protein. Peptides derived from the MAGE-A1 HLA

XX binding protein stimulate the activity and proliferation of CD4+ T

XX lymphocytes. The MAGE-A1 HLA binding protein is useful as a diagnostic

XX agent for diagnosing a disorder characterized by expression of MAGE-A1.

XX The protein is used for treating a disorder characterized by expression

XX of MAGE-A1 such as cancers e.g. melanoma, squamous cell carcinomas,

XX colorectal carcinomas, osteosarcomas, and lymphocytic leukemias.

XX Peptides derived from the MAGE-A1 HLA binding protein are useful in the

XX production of anti-tumour vaccines.

XX Sequence 9 AA;

Query Match 100.0%; Score 59; DB 22; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYQV 9  
 |||||||  
 1 ktwgqyqgv 9

RESULT 27

AAR78643

ID AAR78643 standard; Protein; 10 AA.

XX

AC AAR78643;

XX 22-JAN-1996 (first entry)

XX Immunogenic peptide of melanoma associated antigen gp100.

XX Melanoma: antigen; vaccine; immunogen; primer; probe; detection;

KM Identification; tumour; gp100.

XX Homo sapiens.

XX EP668350-A1.

XX 23-AUG-1995.

XX 14-FEB-1995; 95EP-0200348.

XX 21-DEC-1994; 94EP-0203709.

XX 16-FEB-1994; 94EP-0200337.

PA (ALKU) AKZO NOBEL NV.

XX Adema GJ, Figdor CG;

XX WPI; 1995-284790/38.

XX N-PSDB; AAQ96055.

XX Melanoma associated antigen gp100 - used in vaccines and for the

XX detection of tumours

XX Claim 5; Page 31; 40pp; English.

XX Immunogenic peptides derived from the melanoma associated antigen

XX (see AAR78639-45) may be used in the production of vaccines.

XX Nucleotide sequences encoding the immunogenic peptides may be used

XX as primers and probes in the detection of melanoma cells. Tumour

XX infiltrating lymphocytes capable of binding to the melanoma

XX associated antigen can be cultured ex vivo and returned to melanoma

XX particles, and when radiolabelled, they may be used to identify

XX tumour deposits.

XX Sequence 10 AA;

OY 1 KTWGQYQV 9  
 |||||||  
 1 ktwgqyqgv 9

Db 1 ktwgqyqgv 9

RESULT 28

AAR84209

ID AAR84209 standard; Peptide; 10 AA.

XX AAR84209;

XX 25-APR-1996 (first entry)

XX gp100 melanoma antigen immunogenic peptide (G10-154).

XX gp100; melanoma antigen recognised by T-cells; MART; melanoma;

XX metastatic melanoma; tumour-associated antigen;

XX immunogenic peptide; diagnosis; prognosis; prophylaxis;

XX therapy; vaccine.

XX Synthetic.

XX WO9529193-A2.

XX 02-NOV-1995.

XX 21-APR-1995; 95WO-US05063.  
XX  
XX 05-APR-1995; 95US-0417174.  
PR 22-APR-1994; 94US-0231565.  
XX  
XX (USSH ) US SEC DEPT HEALTH.  
XX  
XX Kawakami Y, Rosenberg SA;  
XX  
XX WPI, 1995-382963/49.  
XX  
XX DNA encoding melanoma antigens recognised by T-lymphocytes - also  
PT vectors, host cells and antibodies, used to detect, treat and  
PT immunise animal against melanoma.  
XX  
XX Claim 55; Page 131; 184pp; English.  
XX  
XX The immunogenic peptide is derived from CDNA25 (AAR84854), a  
CC melanoma antigen derivative of gp100 (see AAR84855). The  
CC peptide and its derivatives (see AAR84200-R84211) are used in  
CC medicaments (vaccines) for the treatment or prevention (by  
CC immunization) of melanoma. Antibodies against melanoma-specific  
CC antigens and its immunogenic peptides may be used in the  
CC detection and isolation of the antigen from a sample, the  
CC detection of which is indicative of a disease state  
CC (melanoma or metastatic melanoma).  
XX  
XX Sequence 10 AA;  
SQ

Query Match 100.0%; Score 59; DB 16; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0024;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYWCY 9  
| | | | | | | |  
1 KtwgqywcY 9  
Db

RESULT 29  
AAR78642  
ID AAR78642 standard; Protein; 11 AA.  
XX  
XX AAR78642;  
AC  
XX  
XX 22-JAN-1996 (first entry)  
DT  
XX  
XX Immunogenic peptide of melanoma associated antigen gp100.  
DE  
XX  
XX Melanoma; antigen; vaccine; immunogen; primer; probe; detection;  
KW identification; tumour; gp100.  
XX  
XX Homo sapiens.  
OS  
XX  
XX EP668350-A1.  
PN  
XX  
XX 23-AUG-1995.  
PD  
XX  
XX 14-FEB-1995; 95EP-0200348.  
PF  
XX  
XX 21-DEC-1994; 94EP-0203709.  
PR  
XX  
XX 16-FEB-1994; 94EP-0200337.  
PR  
XX  
XX (ALKU ) AKZO NOBEL NV.  
PA  
XX  
XX Adema GT, Fijdor CG;  
PI  
XX  
XX WPI, 1995-284790/38.  
DR  
XX  
XX N-PSDB; AAQ96055.  
DR  
XX  
XX Melanoma associated antigen gp100 - used in vaccines and for the  
PT detection of tumours

XX Claim 5; Page 31; 40pp; English.  
PS  
XX  
XX Immunogenic peptides derived from the melanoma associated antigen  
CC (See AAR78639-45) may be used in the production of vaccines.  
CC Nucleotide sequences encoding the immunogenic peptides may be used  
CC as primers and probes in the detection of melanoma cells. Tumour  
CC infiltrating lymphocytes capable of binding to the melanoma  
CC associated antigen can be cultured ex vivo and returned to melanoma  
CC particles, and when radiolabelled, they may be used to identify  
CC tumour deposits.  
XX  
XX Sequence 11 AA;  
SQ

Query Match 100.0%; Score 59; DB 16; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0026;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYWCY 9  
| | | | | | | |  
3 KtwgqywcY 11  
Db

RESULT 30  
AAB61647  
ID AAB61647 standard; peptide; 30 AA.  
XX  
XX AAB61647;  
AC  
XX  
XX 09-APR-2001 (first entry)  
DT  
XX  
XX Gp100 peptide #1.  
DE  
XX  
XX Immunostimulant; vaccine; immune response; parathyroid hormone; antibody;  
KW Gp100; melanoma.  
XX  
XX Unidentified.  
OS  
XX  
XX GB2351905-A.  
PN  
XX  
XX 17-JAN-2001.  
PD  
XX  
XX 13-APR-2000; 2000GB-0008992.  
PF  
XX  
XX 13-APR-1999; 99GB-0008263.  
PR  
XX  
XX (BIND-) BINDING SITE LTD.  
PA  
XX  
XX Bradwell AR, Mead GP;  
PI  
XX  
XX WPI, 2001-140512/15.  
DR  
XX  
XX New antigenic compositions for eliciting improved immune responses,  
PT especially in overcoming tolerance to self-antigens -  
PT  
XX  
XX Disclosure; Page 18; 40pp; English.  
PS  
XX  
XX The present invention relates to a composition for eliciting an immune  
CC response to a target molecule comprising an immunological adjuvant and  
CC one or more peptides. The present sequence is one such peptide. The  
CC peptides are similar to, but different from, a target peptide e.g.  
CC Gp100. Gp100 is a protein associated with melanoma. The present peptide  
CC can elicit the production of antibodies against Gp100, and so can be  
CC used to treat melanoma.  
XX  
XX Sequence 30 AA;  
SQ

Query Match 100.0%; Score 59; DB 22; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.0071;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

. Tue Jul 3 20:10:09 2001

us-09-214-836-9.rag

Page 15

QY 1 KTWGQYNOV 9  
| | | | |  
Db 21 ktwgqywov 29

Search completed: July 3, 2001, 10:23:58  
job time: 1383 sec

7



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw<sub>Model</sub>

Run on: July 3, 2001, 10:23:12 ; Search time 23.85 seconds  
(without alignments)  
28.745 Million cell updates/sec

Title: US-09-214-836-9  
Perfect score: 59  
Sequence: 1 KTWGQYMOV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR-68:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	662	2 I38400	melanoma-associate
2	59	100.0	668	2 A41234	melanocyte-specific
3	55	93.2	626	2 S53871	Pmel 17 protein -
4	45	76.3	264	2 H85861	hypothetical prote
5	45	76.3	335	1 A39862	protein-tyrosine-P
6	45	76.3	549	2 H64982	hypothetical prote
7	43	72.9	209	2 S75029	hypothetical prote
8	41	69.5	277	1 QRECSY	sulfate/chitosulat
9	41	69.5	277	2 H85885	hypothetical prote
10	41	69.5	315	2 T07314	cytochrome c-type
11	40	67.8	281	2 A82104	conserved hypotet
12	40	67.8	305	1 S52775	hypothetical prote
13	40	67.8	400	2 S76446	hypothetical prote
14	40	67.8	807	2 F64844	hypothetical prote
15	40	67.8	807	2 F85647	ycods protein precu
16	39	66.1	208	2 S46301	probable outer mem
17	39	66.1	236	2 A75530	cytochrome c-type
18	39	66.1	276	2 A72451	probable lactose t
19	39	66.1	794	2 S73378	probable lipoprote
20	38	64.4	160	2 T09871	expansin - upland
21	38	64.4	232	2 T09818	expansin (clone pp
22	38	64.4	232	2 T09821	expansin (clone pp
23	38	64.4	232	2 T09825	expansin (clone pp
24	38	64.4	232	2 T09826	expansin (clone pp
25	38	64.4	237	2 T50654	expansin EXP1 (imp
26	38	64.4	246	2 T04175	expansin - rice
27	38	64.4	248	2 C84444	probable expansin
28	38	64.4	250	2 T10079	expansin S1 precu
29	38	64.4	251	2 T03298	expansin 2 - rice

30	38	64.4	253	2 F84831	probable expansin
31	38	64.4	255	2 T06573	expansin 18 - toma
32	38	64.4	255	2 T50655	expansin EXP5 (imp
33	38	64.4	255	2 T50656	expansin EXP2 (imp
34	38	64.4	257	2 T02727	probable expansin
35	38	64.4	257	2 D84820	probable expansin
36	38	64.4	257	2 T50658	expansin 9 (import
37	38	64.4	258	2 T09786	expansin - upland
38	38	64.4	258	2 S53082	pollen allergen ho
39	38	64.4	258	2 T48247	expansin-1-like prot
40	38	64.4	259	2 T50653	expansin EXP6 (imp
41	38	64.4	260	2 T08016	probable expansin
42	38	64.4	260	2 T47689	expansin-like prot
43	38	64.4	261	2 T07630	expansin 1 - tomat
44	38	64.4	262	2 T02530	probable expansin
45	38	64.4	262	2 T50660	alpha-expansin 2 (

#### ALIGNMENTS

RESULT 1  
I38400  
melanoma-associated ME20 antigen (me20m) - human  
N:Alternate names: melanoma antigen 25  
C:Species: Homo sapiens (man)  
C:Date: 01-Nov-1996 #sequence\_revision 01-Nov-1996 #text\_change 01-Dec-2000  
C:Accession: I38400; A53668; A55753  
R:Marsh, G.A.; Marken, J.S.; Neubauer, M.; Aruffo, A.; Hellstrom, I.; Hellstrom, K.;  
DNA Cell Biol. 13, 87-95, 1994  
A:Title: Cloning and expression of the gene for the Melanoma-Associated ME20 Antigen.  
A:Reference number: I38400; MUID:94235165  
A:Accession: I38400  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-662 <RES>  
A:Cross-references: EMBL:U01874; NID:9494939; PIDN:AA18479.1; PID:9494940  
R:Adema, G.J.; de Boer, A.J.; Vogel, A.M.; Loenen, W.A.M.; Figdor, C.G.  
J. Biol. Chem. 269, 20126-20133, 1994  
A:Title: Molecular characterization of the melanocyte lineage-specific antigen gp100.  
A:Reference number: A53668; MUID:94327568  
A:Accession: A53668  
A:Molecule type: mRNA  
A:Residues: 1-592,594-662 <ADE>  
A:Cross-references: GB:S73003; NID:9639589; PIDN:AAC60634.1; PID:9639580  
R:Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.; Sakaguchi, K.; Appella, E.  
Proc. Natl. Acad. Sci. U.S.A. 91, 6458-6462, 1994  
A:Title: Identification of a human melanoma antigen recognized by tumor-infiltrating  
A:Reference number: A55753; MUID:94294401  
A:Accession: A55753  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-161,'F',163-592,594-662 <KAW>  
C:Keywords: glycoprotein

Query Match 100.0% Score 59; DB 2; Length 662;  
Best Local Similarity 100.0% Pred. No. 0.13;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
|||||||  
DB 154 KTWGQYMOV 162

RESULT 2  
A41234  
melanocyte-specific protein Pmel-17 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Jun-1992 #sequence\_revision 19-Jun-1992 #text\_change 30-Sep-1993  
C:Accession: A41234  
R:Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.; Gilbert, D.J.; Jenkins,  
Proc. Natl. Acad. Sci. U.S.A. 88, 9228-9232, 1991

A:Title: A melanocyte-specific gene, Pmel 17, maps near the silver coat color locus on H  
A:Reference number: A41234; MUID:92021023  
A:Accession: A41234  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-668 <KMO>  
A:Cross-references: GB:M77348

Query Match  
Best Local Similarity 100.0%; Score 59; DB 2; Length 668;  
Pred. No. 0.13;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
|||||  
Db 154 KTWGQYMOV 162

RESULT 3  
S53871  
Pmel 17 protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 05-Nov-1999  
C:Accession: S53871  
R:Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.; Chintamani, C.; Bennett, D.; Pick  
Nucleic Acids Res. 23, 154-158, 1995  
A:Title: Mouse silver mutation is caused by a single base insertion in the putative cyto  
A:Reference number: S53871; MUID:95175358  
A:Accession: S53871  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-626 <KMO>  
A:Cross-references: GB:U14133; NID:9887940; PIDN:AAA69538.1; PID:9887941

Query Match  
Best Local Similarity 93.2%; Score 55; DB 2; Length 626;  
Pred. No. 0.47;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
|||||  
Db 154 KTWGQYMOV 162

RESULT 4  
H85861  
hypothetical protein 23480 [Imported] - Escherichia coli (strain O157:H7)  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
C:Accession: H85861  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: H85861  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-264 <STO>  
A:Cross-references: GB:AE005174; NID:q12516559; PIDN:AA657356.1; GSPDB:GN00145; UMGCP:234  
A:Experimental source: strain O157:H7, substrain EDJ933  
C:Genetics:  
A:Gene: 23480

Query Match  
Best Local Similarity 76.3%; Score 45; DB 2; Length 264;  
Pred. No. 6.7;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOYMO 8  
|||||  
Db 125 WGOYMO 130

RESULT 5  
A39862  
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 1 - yeast (Saccharomyces

N:Alternate names: protein D0815, protein YDL230w  
C:Species: Saccharomyces cerevisiae  
C:Date: 30-Dec-1991 #sequence\_revision 08-Mar-1996 #text\_change 21-Jul-2000  
C:Accession: A39862; S67793  
R:Guan, K.; Deschenes, R.J.; Qiu, H.; Dixon, J.E.  
J. Biol. Chem. 266, 12964-12970, 1991  
A:Title: Cloning and expression of a yeast protein tyrosine phosphatase.  
A:Reference number: A39862; MUID:91302512  
A:Accession: A39862  
A:Molecule type: DNA  
A:Residues: 1-335 <GUA>  
A:Cross-references: GB:M64062; NID:q172295; PIDN:AAA34923.1; PID:q172296  
R:Rasmussen, S.W.  
submitted to the Protein Sequence Database, July 1996

A:Reference number: S67778  
A:Accession: S67793  
A:Molecule type: DNA  
A:Residues: 1-335 <RAS>  
A:Cross-references: EMBL:Z74278; NID:q1431387; PIDN:CAA98809.1; PID:q1431388; GSPDB:G  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:PTPI; MIPS:YDL230w  
A:Cross-references: SGD:S0002389; MIPS:YDL230w  
A:Map position: 4L  
C:Superfamily: Saccharomyces protein-tyrosine-phosphatase, nonreceptor type 1; protei  
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphat  
F:52-303/Domain: protein-tyrosine-phosphatase homology <PTP>  
F:252/Active site: Cys (phosphocysteine intermediate) #status predicted  
F:258/Binding site: substrate phosphate (Arg) #status predicted

Query Match  
Best Local Similarity 76.3%; Score 45; DB 1; Length 335;  
Pred. No. 8.4;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQYMOV 9  
|||||  
Db 108 KTWGQYMOV 116

RESULT 6  
H64992  
hypothetical protein b2226 - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 04-Mar-2000  
C:Accession: H64992  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: H64992  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-549 <BLAT>  
A:Cross-references: GB:AE00012; GB:U00096; NID:q1788555; PIDN:AA675286.1; PID:q17885  
A:Experimental source: strain K-12, substrain MG1655  
C:Superfamily: Escherichia coli hypothetical protein b2226

Query Match  
Best Local Similarity 76.3%; Score 45; DB 2; Length 549;  
Pred. No. 13;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOYMO 8  
|||||  
Db 125 WGOYMO 130



## RESULT 7

S75029  
hypothetical protein slr2003 - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000

C:Accession: S75029

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-116, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.

A:Reference number: S74322; MUID:97061201

A:Accession: S75029

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-209 <K&N>

A:Cross-references: EMBL:DP90910; GB:AB001339; NID:q1652956; PIDN:BAAL7891.1; PID:q165297

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Superfamily: *Synechocystis* hypothetical protein slr2003

## Query Match

Best Local Similarity 72.9%; Score 43; DB 2; Length 209;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KTWGOYWO 8

Db 75 KTWGOYWO 82

## RESULT 8

ORECST

sulfate/thiosulfate transport protein cyst - *Escherichia coli*

N:Alternate names: sulfate transport system permease protein cyst

C:Species: *Escherichia coli*

C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Jul-1999

C:Accession: A35402; G65016; B35403

R:Sirko, A.; Hrytlewicz, M.; Hulanicka, D.; Boeck, A.

J. Bacteriol. 172, 3351-3357, 1990

A:Title: Sulfate and thiosulfate transport in *Escherichia coli* K-12: nucleotide sequence

A:Reference number: A35402; MUID:90264334

A:Accession: A35402

A:Molecule type: DNA

A:Residues: 1-277 <STR>

A:Cross-references: GB:M32101; GB:M38050; NID:q145657; PIDN:AAA23637.1; PID:q145659

A:Experimental source: strain K12

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C. A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: G65016

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-277 <BLAT>

A:Cross-references: GB:AE000330; GB:U00056; NID:q1788763; PIDN:AACT5477.1; PID:q1788764

A:Experimental source: strain K-12, substrain MG1655

R:Hrytlewicz, M.; Sirko, A.; Palucha, A.; Boeck, A.; Hulanicka, D.

J. Bacteriol. 172, 3358-3366, 1990

A:Title: Sulfate and thiosulfate transport in *Escherichia coli* K-12: identification of a

A:Reference number: A35403; MUID:90264335

A:Accession: B35403

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-126; F, 128-133 <HR>

C:Comment: This is one of the membrane-associated components of the binding protein-depe

C:Genetics:

A:Gene: *cytS*; *cytS*

A:Map position: 52 min

C:Superfamily: maltose transport protein malG

C:Keywords: binding protein-dependent transport system; inner membrane; membrane protein

Query Match 69.5%; Score 41; DB 1; Length 277;

Best Local Similarity 62.5%; Pred. No. 28;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 2 TWGOYWOV 9

Db 43 SMOATWEV 50

## RESULT 9

B85885  
hypothetical protein *cytS* [imported] - *Escherichia coli* (strain O157:H7)

C:Species: *Escherichia coli*

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001

C:Accession: B85885

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimantanta, E.; Potamousis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A:Reference number: A83480; MUID:21074935; PMID:11206531

A:Accession: B85885

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-277 <STO>

A:Cross-references: GB:AE005174; NID:q12516799; PIDN:AA657542.1; GSPDB:GN00145; UMGF:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: *cytS*

C:Superfamily: maltose transport protein malG

Query Match 69.5%; Score 41; DB 2; Length 277;

Best Local Similarity 62.5%; Pred. No. 28;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 2 TWGOYWOV 9

Db 43 SMOATWEV 50

## RESULT 10

T07314  
cytochrome c-type synthesis protein homolog - *Chlorella vulgaris* chloroplast

C:Species: chloroplast *Chlorella vulgaris*

C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000

C:Accession: T07314

R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Na

Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997

A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga *C*

A:Reference number: Z15985; MUID:97303241

A:Accession: T07314

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-315 <W&K>

A:Cross-references: EMBL:AB001684; NID:q2224352; PIDN:BA57962.1; PID:q2224478

C:Genetics:

A:Gene: *ycf5*

A:Superfamily: cytochrome c-type synthesis protein

C:Keywords: chloroplast

Query Match 69.5%; Score 41; DB 2; Length 315;

Best Local Similarity 71.4%; Pred. No. 31;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 KTWGOYWO 7

Db 243 ETWGNW 249

## RESULT 11

AB2104  
conserved hypothetical protein VC2229 [imported] - *Vibrio cholerae* (strain N16961 serogroup C)  
C:Species: *Vibrio cholerae*  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: AB2104  
R:Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Chardson, D.; Esmolaeva, M.D.; Yamathayan, J.; Bass, S.; Qin, H.; Dragoli, I.; Sellers, F. I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A:Reference number: AB2035; MUID:2040933  
A:Accession: AB2104  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-281 <HEI>  
A:Cross-references: GB:AE004294; GB:AE003852; NID:9656774; PIDN:NAF95373.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC2229  
A:Map position: 1  
C:Superfamily: hypothetical protein H11037

Query Match 67.8%; Score 40; DB 2; Length 281;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOYW 7  
|||||  
DB 99 WGOYW 103

RESULT 12  
S52775  
hypothetical protein 2 - *Chloroflexus aurantiacus*  
C:Species: *Chloroflexus aurantiacus*  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: S52775  
R:Niedermeier, G.; Shiozawa, J.A.; Lotspeltch, F.; Feick, R.G.  
FEBS Lett. 342, 61-65, 1994  
A:Title: The primary structure of two chlorosome proteins from *Chloroflexus aurantiacus*.  
A:Reference number: S43678; MUID:94192803  
A:Accession: S52775  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-305 <NIB>  
A:Cross-references: EMBL:Z34000; NID:9496485; PIDN:CA83969.1; PID:9496488  
A:Note: only a part of the coding sequence is given in this paper  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1994  
C:Superfamily: Methanococcus jannaschii conserved hypothetical protein M0279

Query Match 67.8%; Score 40; DB 1; Length 305;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOYW 7  
|||||  
DB 246 WGOYW 250

RESULT 13  
S76446  
hypothetical protein - *Synechocystis* sp. (strain PCC 6803)  
C:Species: *Synechocystis* sp.  
A:Variety: PCC 6803  
C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S76446  
R:Kameko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.

A:Reference number: S74322; MUID:97061201  
A:Accession: S76446  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-400 <KAN>  
A:Cross-references: EMBL:D90915; GB:AB001339; NID:91653604; PIDN:BA18575.1; PID:d101  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 67.8%; Score 40; DB 2; Length 400;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOYW 7  
|||||  
DB 335 WGOYW 339

RESULT 14  
F64844  
ycds protein precursor - *Escherichia coli*  
C:Species: *Escherichia coli*  
C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 04-Mar-2000  
C:Accession: F64844  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of *Escherichia coli* K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: F64844  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-807 <BLAT>  
A:Cross-references: GB:AE000204; GB:U00096; NID:91787256; PIDN:AACT4109.1; PID:917872  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: ycds  
C:Superfamily: *Escherichia coli* ycds protein  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-807/Product: ycds protein #status predicted <MAT>

Query Match 67.8%; Score 40; DB 2; Length 807;  
Best Local Similarity 100.0%; Pred. No. 11e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOYW 7  
|||||  
DB 314 WGOYW 318

RESULT 15  
F85647  
probable outer membrane protein ycds [imported] - *Escherichia coli* (strain O157:H7)  
C:Species: *Escherichia coli*  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
C:Accession: F85647  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
Miller, L.; Grobeck, E.J.; Davis, N.W.; Llm, A.; Dimalanta, E.; Potamousis, K.; Apoda  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
A:Reference number: AB5480; MUID:21074935; PMID:11206551  
A:Accession: F85647  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-807 <STO>  
A:Cross-references: GB:AE005174; NID:912514389; PIDN:AGC5642.1; GSPDB:GN00145; OMGP:  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: ycds  
C:Superfamily: *Escherichia coli* ycds protein

Query Match 67.8%; Score 40; DB 2; Length 807;  
Best Local Similarity 100.0%; Pred. No. 1;le+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TWGQY 7  
|||||  
Db 314 TWGQY 318

RESULT 16  
S46301

fucoxanthin chlorophyll a/c-binding light-harvesting protein - Isochrysis galbana

N:Alternate names: RCP protein

C:Species: Isochrysis galbana

C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 05-Nov-1999

C:Accession: S46301

R:LaRoche, J.; Henry, D.; Wyman, K.; Sukenik, A.; Falkowski, P.

Plant Mol. Biol. 25, 355-368, 1994

A:Title: Cloning and nucleotide sequence of a cDNA encoding a major fucoxanthin-, chlorophyll family.

A:Reference number: S46301; MUID:94325461

A:Accession: S46301

A:Molecule type: mRNA

A:Residues: 1-208 <IAR>

A:Cross-references: EMBL:X77333; NID:9535080; PIDN:CA54547.1; PID:9535081

A:Experimental source: cultivar DJN

C:Keywords: light-harvesting complex

Query Match 66.1%; Score 39; DB 2; Length 208;  
Best Local Similarity 83.3%; Pred. No. 42;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 TWGQY 7  
|||||  
Db 198 TWGQY 203

RESULT 17  
A75530

cytochrome c-type biogenesis protein, heme exporter protein C - Deinococcus radiodurans

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000

C:Accession: A75530

R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: A75530

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-236 <WHI>

A:Cross-references: GB:AE001895; GB:AE000513; NID:96458024; PIDN:AF09930.1; PID:9645802

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0348

A:Map position: 1

C:Superfamily: helic protein

Query Match 66.1%; Score 39; DB 2; Length 236;  
Best Local Similarity 83.3%; Pred. No. 48;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 TWGQY 7  
|||||  
Db 101 TWGQY 106

RESULT 18  
A72451

Probable lactose transport system permease protein APE2253 - Aeropyrum pernix (strain C:Species: Aeropyrum pernix  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: A72451  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.

DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A:Reference number: A72450; MUID:99310339

A:Accession: A72451

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-276 <KAW>

A:Cross-references: DDBJ:AP000064; NID:95105945; PIDN:BA81265.1; PID:95105954

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2253

C:Superfamily: inner membrane protein upga

Query Match 66.1%; Score 39; DB 2; Length 276;  
Best Local Similarity 55.6%; Pred. No. 55;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGQY 9  
|||||  
Db 215 KTWGQY 223

RESULT 19  
S73328

probable lipoprotein E07orf794 - Mycoplasma pneumoniae (strain ATCC 29342)

N:Alternate names: MG260 homolog E07orf794

C:Species: Mycoplasma pneumoniae

A:Variety: ATCC 29342

C>Date: 26-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-Dec-1999

C:Accession: S73328

R:Himmelfreih, R.; Hilbert, H.; Plagens, H.; Plöckl, E.; Li, B.C.; Hermann, R.

Nucleic Acids Res. 24, 4420-4449, 1996

A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumon

A:Reference number: S73327; MUID:97105885

A:Accession: S73328

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-794 <HIM>

A:Cross-references: EMBL:AE000001; GB:U00089; NID:91673645; PIDN:AB95650.1; PID:9167

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996

C:Genetics:

A:Genetic code: SGC3

C:Superfamily: hypothetical protein MG185

Query Match 66.1%; Score 39; DB 2; Length 794;  
Best Local Similarity 66.7%; Pred. No. 1;5e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 KTWGQY 9  
|||||  
Db 268 KTWGQY 276

RESULT 20  
T09871

expansin - upland cotton (fragment)

C:Species: Gossypium hirsutum (upland cotton)

C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000

C:Accession: T09871

R:Shimizu, Y.; Aotsuka, S.; Hasegawa, O.; Kawada, T.; Sakuno, T.; Sakai, F.; Hayashi,

Plant Cell Physiol. 38, 373-376, 1997

A:Title: Changes in levels of mRNAs for cell wall-related enzymes in growing cotton f

A:Reference number: Z16889; MUID:97294938

A:Accession: T09871

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA  
 A:Residues: 1-160 <SH1>  
 A:Cross-references: EMBL:D88415; NID:g2244735; PIDN:BAA21109.1; PID:g2244736  
 A:Experimental source: strain Coker12; fiber  
 C:Superfamily: expansin

Query Match 64.4%; Score 38; DB 2; Length 160;  
 Best Local Similarity 62.5%; Pred. No. 47;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 KTWGQYWO 8  
 Db 110 RNMGQNMW 117

RESULT 21  
 T09818  
 expansin (clone ptxp2) - lobliolly pine (fragment)  
 C:Species: Pinus taeda (lobliolly pine)  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 26-May-2000  
 C:Accession: T09818  
 R:Hutchinson, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.  
 submitted to the EMBL Data Library, July 1996  
 A:Description: Expansins are conserved in conifers and expressed in response to exogenous  
 A:Reference number: Z16866  
 A:Accession: T09818  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-232 <HUT>  
 A:Cross-references: EMBL:U64890; NID:g1778098; PID:g1778099  
 A:Experimental source: clone ptxp2; hypocotyl  
 C:Superfamily: expansin

Query Match 64.4%; Score 38; DB 2; Length 232;  
 Best Local Similarity 62.5%; Pred. No. 67;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 KTWGQYWO 8  
 Db 182 RNMGQNMW 189

RESULT 22  
 T09821  
 expansin (clone ptxp3) - lobliolly pine (fragment)  
 C:Species: Pinus taeda (lobliolly pine)  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 26-May-2000  
 C:Accession: T09821  
 R:Hutchinson, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.  
 submitted to the EMBL Data Library, July 1996  
 A:Description: Expansins are conserved in conifers and expressed in response to exogenous  
 A:Reference number: Z16866  
 A:Accession: T09821  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-232 <HUT>  
 A:Cross-references: EMBL:U64891; NID:g1778100; PID:g1778101  
 A:Experimental source: clone ptxp3  
 C:Superfamily: expansin

Query Match 64.4%; Score 38; DB 2; Length 232;  
 Best Local Similarity 62.5%; Pred. No. 67;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 KTWGQYWO 8  
 Db 182 RNMGQNMW 189

RESULT 23

T09825  
 expansin (clone ptxp4) - lobliolly pine (fragment)  
 C:Species: Pinus taeda (lobliolly pine)  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 26-May-2000  
 C:Accession: T09825  
 R:Hutchinson, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.  
 submitted to the EMBL Data Library, July 1996  
 A:Description: Expansins are conserved in conifers and expressed in response to exogenous  
 A:Reference number: Z16866  
 A:Accession: T09825  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-232 <HUT>  
 A:Cross-references: EMBL:U64892; NID:g1778102; PID:g1778103  
 A:Experimental source: clone ptxp4  
 C:Superfamily: expansin

Query Match 64.4%; Score 38; DB 2; Length 232;  
 Best Local Similarity 62.5%; Pred. No. 67;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 KTWGQYWO 8  
 Db 182 RNMGQNMW 189

RESULT 24  
 T09826  
 expansin (clone ptxp5) - lobliolly pine (fragment)  
 C:Species: Pinus taeda (lobliolly pine)  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 26-May-2000  
 C:Accession: T09826  
 R:Hutchinson, K.W.; Singer, P.B.; Diaz-Sala, C.; Greenwood, M.S.  
 submitted to the EMBL Data Library, July 1996  
 A:Description: Expansins are conserved in conifers and expressed in response to exogenous  
 A:Reference number: Z16866  
 A:Accession: T09826  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-232 <HUT>  
 A:Cross-references: EMBL:U64893; NID:g1778104; PID:g1778105  
 A:Experimental source: clone ptxp5  
 C:Superfamily: expansin

Query Match 64.4%; Score 38; DB 2; Length 232;  
 Best Local Similarity 62.5%; Pred. No. 67;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 KTWGQYWO 8  
 Db 182 RNMGQNMW 189

RESULT 25  
 T50654  
 expansin EXP1 [imported] - Arabidopsis thaliana (fragment)  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 02-Sep-2000  
 C:Accession: T50654  
 R:Shcherban, T.Y.; Shi, J.; Durachko, D.M.; Guiltinan, M.J.; McQueen-Mason, S.J.; Shi  
 Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995  
 A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, mu  
 A:Reference number: Z14894; MUID:96016146  
 A:Accession: T50654  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-237 <SHC>  
 A:Cross-references: EMBL:U30476; PIDN:AAB38070.1  
 C:Genetics:  
 A:Gene: EXP1  
 C:Function:

A:Description: induces extension (creep) in plant cell walls  
C:Superfamily: expansin  
C:Keywords: cell wall

Query Match 64.4%; Score 38; DB 2; Length 237;  
Best Local Similarity 62.5%; Pred. No. 68;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGQYMQ 8  
: ||| ||  
Db 186 RNMGQNMQ 193

RESULT 26  
T04175  
expansin - rice  
C:Species: Oryza sativa (rice)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 26-May-2000  
C:Accession: T04175  
R:Cho, H.T.; Kende, H.  
Plant Cell 9, 1661-1671, 1997  
A:Title: Expression of expansin genes is correlated with growth in deepwater ce.  
A:Reference number: 215042; MUID:97480100  
A:Accession: T04175  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-246 <CHO>  
A:Cross-references: EMBL:U085246; NID:G1815680; PIDN:AMB81662.1; PID:G1815681  
A:Experimental source: cv. Pin Gaew 56  
C:Genetics:  
A:Gene: EXP4  
C:Superfamily: expansin

Query Match 64.4%; Score 38; DB 2; Length 246;  
Best Local Similarity 62.5%; Pred. No. 70;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGQYMQ 8  
: ||| ||  
Db 196 RNMGQNMQ 203

RESULT 27  
C84444  
probable expansin [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: C84444  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffit, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
Euss, D.; Nleman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Ventner, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487  
A:Accession: C84444  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-248 <STO>  
A:Cross-references: GB:AE002093; NID:G3461833; PIDN:AMC32927.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g03090  
A:Map position: 2  
C:Superfamily: expansin

Query Match 64.4%; Score 38; DB 2; Length 248;  
Best Local Similarity 62.5%; Pred. No. 71;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGQYMQ 8  
: ||| ||

Db 197 RNMGQNMQ 204

RESULT 28  
T10079  
expansin S1 precursor - cucumber  
C:Species: Cucumis sativus (cucumber)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T10079  
R:Shcherban, T.Y.; Shi, J.; Durachko, D.M.; Gullitman, M.J.; McQueen-Mason, S.J.; Shi  
Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995  
A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, mu  
A:Reference number: 214894; MUID:96016146  
A:Accession: T10079  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-250 <SHC>  
A:Cross-references: EMBL:U030382; NID:G1040874; PIDN:AMB37746.1; PID:G1040875  
A:Experimental source: cultivar Burpee Pickler  
C:Genetics:  
A:Gene: EXP1  
C:Function:  
A:Description: mediates cell wall extension  
C:Superfamily: expansin  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-250/Product: expansin #status predicted <MAT>

Query Match 64.4%; Score 38; DB 2; Length 250;  
Best Local Similarity 62.5%; Pred. No. 71;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGQYMQ 8  
: ||| ||  
Db 200 RNMGQNMQ 207

RESULT 29  
T03298  
expansin 2 - rice  
C:Species: Oryza sativa (rice)  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 21-Jul-2000  
C:Accession: T03298  
R:Shcherban, T.Y.; Shi, J.; Durachko, D.M.; Gullitman, M.J.; McQueen-Mason, S.J.; Shi  
Proc. Natl. Acad. Sci. U.S.A. 92, 9245-9249, 1995  
A:Title: Molecular cloning and sequence analysis of expansins--a highly conserved, mu  
A:Reference number: 214894; MUID:96016146  
A:Accession: T03298  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-251 <SHC>  
A:Cross-references: EMBL:U030477; NID:G1041709; PIDN:AMB38074.1; PID:G1041710  
C:Genetics:  
A:Gene: EXP2  
C:Function:  
A:Description: induces extension (creep) in plant cell  
C:Superfamily: expansin

Query Match 64.4%; Score 38; DB 2; Length 251;  
Best Local Similarity 62.5%; Pred. No. 72;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KTWGQYMQ 8  
: ||| ||  
Db 201 RNMGQNMQ 208

RESULT 30  
F84831  
probable expansin [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001

C:Accession: F84831  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
Evens, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, A.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: A84420; MUID:20083487  
A:Accession: F84831  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-253 <STO>  
A:Cross-references: GB:AE002093; NID:g2651297; PIDN:AAB87577.1; GSPDB:GN00139  
C:genetics:  
A:gene: Atg40610  
A:Map position: 2  
C:Superfamily: expansin

Query Match 64.48; Score 38; DB 2; Length 253;  
Best Local Similarity 62.58; Pred. No. 72;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 KTWGQYWG 8  
Db 203 RNWGQNWQ 210

Search completed: July 3, 2001, 10:23:12  
Job time: 1337 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:22:35 ; Search time 21.69 Seconds  
(without alignments)  
8.359 Million cell updates/sec

Title: US-09-214-836-9

Sequence: 1 KTWGQYMQV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCBUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	9	1 US-08-787-547-61	Sequence 61, App1
2	59	100.0	9	2 US-08-417-174-46	Sequence 46, App1
3	59	100.0	9	2 US-08-902-516-29	Sequence 29, App1
4	59	100.0	9	2 US-09-036-582-26	Sequence 26, App1
5	59	100.0	10	2 US-08-417-174-47	Sequence 47, App1
6	59	100.0	661	2 US-08-417-174-121	Sequence 121, App1
7	59	100.0	668	1 US-07-891-9426-6	Sequence 6, App1
8	55	93.2	661	2 US-08-417-174-27	Sequence 27, App1
9	55	93.2	661	2 US-08-231-565A-27	Sequence 27, App1
10	55	93.2	661	2 US-09-007-961-27	Sequence 27, App1
11	54	91.5	9	2 US-08-417-174-71	Sequence 71, App1
12	54	91.5	9	2 US-08-417-174-72	Sequence 72, App1
13	54	91.5	9	2 US-08-417-174-73	Sequence 73, App1
14	54	91.5	9	2 US-08-417-174-74	Sequence 74, App1
15	54	91.5	9	2 US-08-417-174-75	Sequence 75, App1
16	53	89.8	9	2 US-08-417-174-68	Sequence 68, App1
17	53	89.8	9	2 US-08-417-174-69	Sequence 69, App1
18	53	89.8	9	2 US-08-417-174-70	Sequence 70, App1
19	50	84.7	9	2 US-08-417-174-76	Sequence 76, App1
20	49	83.1	9	2 US-08-417-174-77	Sequence 77, App1
21	49	83.1	9	2 US-08-417-174-78	Sequence 78, App1
22	49	83.1	9	2 US-08-417-174-79	Sequence 79, App1
23	49	83.1	9	2 US-08-417-174-80	Sequence 80, App1
24	49	83.1	9	2 US-08-417-174-81	Sequence 81, App1
25	49	83.1	9	2 US-08-417-174-82	Sequence 82, App1
26	38	64.4	12	1 US-08-439-817-205	Sequence 205, App1
27	38	64.4	20	1 US-08-241-054-85	Sequence 85, App1

28	38	64.4	20	1	US-08-241-054-91	Sequence 91, App1
29	38	64.4	20	1	US-08-390-156A-40	Sequence 40, App1
30	38	64.4	20	1	US-08-390-156A-43	Sequence 43, App1
31	38	64.4	20	1	US-08-439-817-65	Sequence 65, App1
32	38	64.4	20	1	US-08-439-817-71	Sequence 71, App1
33	38	64.4	20	1	US-08-485-508-85	Sequence 85, App1
34	38	64.4	20	1	US-08-485-508-91	Sequence 91, App1
35	38	64.4	72	1	US-08-482-782B-6	Sequence 6, App1
36	38	64.4	72	1	US-08-486-036A-6	Sequence 6, App1
37	38	64.4	72	5	PCT-US96-09848-6	Sequence 6, App1
38	38	64.4	167	2	US-08-845-539-4	Sequence 4, App1
39	38	64.4	179	2	US-08-845-539-6	Sequence 6, App1
40	38	64.4	222	2	US-08-440-517A-3	Sequence 3, App1
41	38	64.4	225	2	US-08-845-539-2	Sequence 2, App1
42	38	64.4	225	2	US-08-440-517A-5	Sequence 5, App1
43	38	64.4	226	2	US-08-440-517A-6	Sequence 6, App1
44	38	64.4	228	2	US-08-440-517A-2	Sequence 2, App1
45	38	64.4	401	5	PCT-US96-09848-31	Sequence 31, App1

#### ALIGNMENTS

RESULT 1  
US-08-787-547-61  
Sequence 61, Application US/08787547  
Patent No. 5783567  
GENERAL INFORMATION:  
APPLICANT: Hedley, Mary Lynne  
APPLICANT: Curley, Joanne M.  
APPLICANT: Langer, Robert S.  
TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY  
TITLE OF INVENTION: OF NOCLEIC ACID  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/787,547  
FILING DATE: 22-JAN-1997  
CLASSIFICATION: 51A  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 08191/003001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-787-547-61  
Query Match 100.0%; Score 59; DB 1; Length 9;  
Best local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGOYMOV 9  
|11111111|  
Db 1 KTWGOYMOV 9

## RESULT 2

US-08-417-174-46  
; Sequence 46, Application US/08417174  
; Patent No. 5844075  
; GENERAL INFORMATION:  
; APPLICANT: KAKAKAMI, YUTAKA; ROSENBERG,  
; APPLICANT: STEVEN A.  
; TITLE OF INVENTION: MELANOMA ANTIGENS AND  
; THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
; TITLE OF INVENTION: METHODS  
; NUMBER OF SEQUENCES: 126  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/417,174  
; FILING DATE: 05-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/231,565  
; FILING DATE: 22-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAROL M. GRUPTI  
; REGISTRATION NUMBER: 37,341  
; REFERENCE/DOCKET NUMBER: 2026-4124US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; TELEX: 421792  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: Peptide  
; US-08-417-174-46

Query Match 100.0%; Score 59; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGOYMOV 9  
|11111111|  
Db 1 KTWGOYMOV 9

RESULT 3  
US-08-902-516-29  
; Sequence 29, Application US/08902516  
; Patent No. 5891432  
; GENERAL INFORMATION:  
; APPLICANT: SOO HOO, WILLIAM  
; TITLE OF INVENTION: MEMBRANE-BOUND CYTOKINE COMPOSITIONS  
; TITLE OF INVENTION: COMPRISING GM-CSF AND METHODS OF MODULATING AN IMMUNE  
; TITLE OF INVENTION: RESPONSE USING SAME  
; NUMBER OF SEQUENCES: 50

## CORRESPONDENCE ADDRESS:

ADDRESSEE: CAMPBELL & FLORES, LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States  
ZIP: 92121  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/902,516  
; FILING DATE: 29-JUL-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-IM 2442  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619)535-9001  
; TELEFAX: (619)535-8949  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-902-516-29

Query Match 100.0%; Score 59; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGOYMOV 9  
|11111111|  
Db 1 KTWGOYMOV 9

## RESULT 4

US-09-036-582-26  
; Sequence 26, Application US/09036582A  
; Patent No. 5965381  
; GENERAL INFORMATION:  
; APPLICANT: van der Bruggen, Pierre  
; TITLE OF INVENTION: DELIVERY OF PROTEINS INTO EUKARYOTIC CELLS.  
; FILE REFERENCE: 11154  
; CURRENT APPLICATION NUMBER: US/09/036,582A  
; CURRENT FILING DATE: 1998-03-06  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 26  
; LENGTH: 9  
; TYPE: PPT  
; ORGANISM: Human gp100Pmel117 peptide  
; US-09-036-582-26

Query Match 100.0%; Score 59; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KTWGOYMOV 9  
|11111111|  
Db 1 KTWGOYMOV 9

RESULT 5  
US-08-417-174-47



Sequence 47, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 751-6849  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-47

Query Match 100.0%; Score 59; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGQYMOV 9  
Db 1 KTWGQYMOV 9

RESULT 6  
US-08-417-174-121  
Sequence 121, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 751-6849  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 121:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Protein  
US-08-417-174-121

Query Match 100.0%; Score 59; DB 2; Length 661;  
Best Local Similarity 100.0%; Pred. No. 0.097;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGQYMOV 9  
Db 154 KTWGQYMOV 162

RESULT 7  
US-07-891-9426-6  
Sequence 6, Application US/07891942G  
Patent No. 5679511  
GENERAL INFORMATION:  
APPLICANT: KWON, BYOUNG SE  
TITLE OF INVENTION: CDNA CLONES FOR HUMAN TYROSINASE AND FOR  
TITLE OF INVENTION: A REGULATORY PROTEIN IN THE MELANIN PROTEIN PATHWAY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Christopher A. Michaels, Barnard, Brown &  
STREET: 306 East State Street; Suite 220  
CITY: Ithaca  
STATE: NY  
COUNTRY: USA  
ZIP: 14850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/891,942G  
FILING DATE: 01-JUN-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 06/915,753  
FILING DATE: 06-OCT-1986  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/362,847  
FILING DATE: 07-JUN-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Michaels, Christopher A

REGISTRATION NUMBER: 34,390  
REFERENCE/DOCKET NUMBER: INDI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 607-273-1711  
TELEFAX: 607-273-2609  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 668 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PUBLICATION INFORMATION:  
AUTHORS: Kwon, Byoung Se  
AUTHORS: Chintamaneni, Chaya  
AUTHORS: Kozak, Christine A  
AUTHORS: Copeland, Neal G  
AUTHORS: Gilbert, Debra J  
AUTHORS: Jenkins, Nancy  
AUTHORS: Barton, David  
AUTHORS: Francke, Uta  
AUTHORS: Kobayashi, Yvonne  
AUTHORS: Kim, Kack K  
TITLE: A melanocyte-specific gene, Pmel 17, maps  
TITLE: near the silver coat color locus on mouse  
TITLE: chromosome 10 and is in a syntenic region on human  
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.  
VOLUME: 88  
PAGES: 9228-9232  
DATE: October-1991  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 668  
US-07-891-942G-6

Query Match 100.0%; Score 59; DB 1; Length 668;  
Best Local Similarity 100.0%; Pred. No. 0.098;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGQYMW 9  
DB 154 KTWGQYMW 162

RESULT 8  
US-08-417-174-27  
Sequence 27, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAMAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Protein  
US-08-417-174-27

Query Match 93.2%; Score 55; DB 2; Length 661;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGQYMW 8  
DB 154 KTWGQYMW 161

RESULT 9  
US-08-231-565A-27  
Sequence 27, Application US/08231565A  
Patent No. 5874560  
GENERAL INFORMATION:  
APPLICANT: KAMAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565A  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Protein  
US-08-231-565A-27

Query Match 93.2%; Score 55; DB 2; Length 661;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGQYWO 8  
Db 154 KTWGQYWO 161

RESULT 10

US-09-007-961-27  
Sequence 27, Application US/09007961  
Patent No. 5994523

GENERAL INFORMATION:

APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,

APPLICANT: STEVEN A.

TITLE OF INVENTION: MELANOMA ANTIGENS AND

TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC

TITLE OF INVENTION: METHODS

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/007, 961

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/231,565

FILING DATE: 22-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: CAROL M. GRUPPI

REGISTRATION NUMBER: 37,341

REFERENCE/DOCKET NUMBER: 2026-4124

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

TELEX: 421792

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 661

TYPE: amino acid

STRANDEDNESS: Unknown

TOPOLOGY: Unknown

MOLECULE TYPE: Protein

US-09-007-961-27

Query Match 93.2%; Score 55; DB 2; Length 661;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGQYWO 8  
Db 154 KTWGQYWO 161

RESULT 11

US-08-417-174-71

Sequence 71, Application US/08417174

Patent No. 5844075

GENERAL INFORMATION:

APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,

APPLICANT: STEVEN A.

TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 126

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/417,174

FILING DATE: 05-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/231,565

FILING DATE: 22-APR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CAROL M. GRUPPI

REGISTRATION NUMBER: 37,341

REFERENCE/DOCKET NUMBER: 2026-4124US1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

TELEX: 421792

INFORMATION FOR SEQ ID NO: 71:

SEQUENCE CHARACTERISTICS:

LENGTH: 9

TYPE: amino acid

STRANDEDNESS: Unknown

TOPOLOGY: Unknown

MOLECULE TYPE: Peptide

US-08-417-174-71

Query Match 91.5%; Score 54; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TWGQYWO 9  
Db 2 TWGQYWO 9

RESULT 12

US-08-417-174-72

Sequence 72, Application US/08417174

Patent No. 5844075

GENERAL INFORMATION:

APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,

APPLICANT: STEVEN A.

TITLE OF INVENTION: MELANOMA ANTIGENS AND

TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC

TITLE OF INVENTION: METHODS

NUMBER OF SEQUENCES: 126

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151  
2152  
2153  
2154  
2155  
2156  
2157  
2158  
2159  
2160  
2161  
2162  
2163  
2164  
2165  
2166  
2167  
2168  
2169  
2170  
2171  
2172  
2173  
2174  
2175  
2176  
2177  
2178  
2179  
2180  
2181  
2182  
2183  
2184  
2185  
2186  
2187  
2188  
2189  
2190  
2191  
2192  
2193  
2194  
2195  
2196  
2197  
2198  
2199  
2200  
2201  
2202  
2203  
2204  
2205  
2206  
2207  
2208  
2209  
2210  
2211  
2212  
2213  
2214  
2215  
2216  
2217  
2218  
2219  
2220  
2221  
2222  
222

RESULT 15  
US-08-417-174-75  
Sequence 75, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 75:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-75

Query Match 91.5%; Score 54; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TWGQYMOV 9  
DB 2 TWGQYMOV 9

RESULT 16  
US-08-417-174-68  
Sequence 68, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE

CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-68

Query Match 89.8%; Score 53; DB 2; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGQYMOV 9  
DB 1 KTWGQYMOV 9

RESULT 17  
US-08-417-174-69  
Sequence 69, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
TITLE OF INVENTION: METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435

Same as  
w/09/24/93  
240995  
Seq No 68

ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-69

Query Match 89.8%; Score 53; DB 2; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGOYMOV 9  
1 |||||  
Db 1 KMWGOYMOV 9

RESULT 18  
US-08-417-174-70  
Sequence 70, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
NUMBER OF INVENTIONS: METHODS  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-70

Query Match 89.8%; Score 53; DB 2; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.4e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTWGOYMOV 9  
1 |||||  
Db 1 KMWGOYMOV 9

RESULT 19  
US-08-417-174-76  
Sequence 76, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
NUMBER OF INVENTIONS: METHODS  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-76

Query Match 84.7%; Score 50; DB 2; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.4e+05;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTWGOYMOV 9  
1 |||||  
Db 1 KTWGOYMOV 9

RESULT 20  
US-08-417-174-77  
Sequence 77, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:

APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 77:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-77

Query Match 83.1%; Score 49; DB 2; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.4e+05;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTMGOYMOV 9  
Db 1 KTMGOYMOV 9

RESULT 21  
US-08-417-174-78  
Sequence 78, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-78

Query Match 83.1%; Score 49; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGOYMOV 9  
Db 3 WGOYMOV 9

RESULT 22  
US-08-417-174-79  
Sequence 79, Application US/08417174  
Patent No. 5844075  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849

TELEX: 421792  
INFORMATION FOR SEQ ID NO: 79;  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-79

Query Match 83.1%; Score 49; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOYMOV 9  
|||||||  
DB 3 WGOYMOV 9

RESULT 23

US-08-417-174-80  
Sequence 80, Application US/08417174  
Patent No. 5844075

GENERAL INFORMATION:

APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.

TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC

TITLE OF INVENTION: METHODS

NUMBER OF SEQUENCES: 126

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/417,174

FILING DATE: 05-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/231,565

FILING DATE: 22-APR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CAROL M. GRUPPI

REGISTRATION NUMBER: 37,341

REFERENCE/DOCKET NUMBER: 2026-4124US1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

TELEX: 421792

INFORMATION FOR SEQ ID NO: 80;

SEQUENCE CHARACTERISTICS:

LENGTH: 9

TYPE: amino acid

STRANDEDNESS: Unknown

TOPOLOGY: Unknown

MOLECULE TYPE: Peptide

US-08-417-174-80

Query Match 83.1%; Score 49; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOYMOV 9

DB 3 WGOYMOV 9  
|||||||

RESULT 24

US-08-417-174-81

Sequence 81, Application US/08417174

Patent No. 5844075

GENERAL INFORMATION:

APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.

TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC

TITLE OF INVENTION: METHODS

NUMBER OF SEQUENCES: 126

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/417,174

FILING DATE: 05-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/231,565

FILING DATE: 22-APR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CAROL M. GRUPPI

REGISTRATION NUMBER: 37,341

REFERENCE/DOCKET NUMBER: 2026-4124US1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

TELEX: 421792

INFORMATION FOR SEQ ID NO: 81;

SEQUENCE CHARACTERISTICS:

LENGTH: 9

TYPE: amino acid

STRANDEDNESS: Unknown

TOPOLOGY: Unknown

MOLECULE TYPE: Peptide

US-08-417-174-81

Query Match 83.1%; Score 49; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOYMOV 9  
|||||||  
DB 3 WGOYMOV 9

RESULT 25

US-08-417-174-82

Sequence 82, Application US/08417174

Patent No. 5844075

GENERAL INFORMATION:

APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.

TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC

TITLE OF INVENTION: METHODS

NUMBER OF SEQUENCES: 126

CORRESPONDENCE ADDRESS:



ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION NUMBER: US/08/417,174  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124051  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 82:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Peptide  
US-08-417-174-82

Query Match 83.1%; Score 49; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.4e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 WGOYMOV 9  
|||||  
Db 3 WGOYMOV 9

RESULT 26  
US-08-439-817-205  
Sequence 205, Application US/08439817  
Patent No. 5728802  
GENERAL INFORMATION:  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Cwirla, Steven E.  
APPLICANT: Dower, William J.  
APPLICANT: Koller, Kerry J.  
APPLICANT: Lee, Jung  
APPLICANT: Martens, Christine L.  
APPLICANT: Ruhland-Fritsch, Beatrice  
TITLE OF INVENTION: Peptides and Compounds That Bind  
TITLE OF INVENTION: Selectins Including Endothelial Leukocyte Adhesion  
NUMBER OF SEQUENCES: 209  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Affymax Technologies, NV  
STREET: 4001 Miranda Ave.  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/439,817  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/241,054  
FILING DATE: 11-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/057,295  
FILING DATE: 05-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/881,395  
FILING DATE: 06-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens, Lauren L.  
REGISTRATION NUMBER: 36,691  
REFERENCE/DOCKET NUMBER: 000324-046/1056.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-496-2300  
TELEFAX: 415-424-0832  
INFORMATION FOR SEQ ID NO: 205:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-439-817-205

Query Match 64.4%; Score 38; DB 1; Length 12;  
Best Local Similarity 62.5%; Pred. No. 2.6;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TWGQLMDL 10  
|||||  
Db 3 TWGQLMDL 10

RESULT 27  
US-08-241-054-85  
Sequence 85, Application US/08241054  
Patent No. 5643873  
GENERAL INFORMATION:  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Cwirla, Steven E.  
APPLICANT: Dower, William J.  
APPLICANT: Koller, Kerry J.  
APPLICANT: Lee, Jung  
APPLICANT: Martens, Christine L.  
APPLICANT: Ruhland-Fritsch, Beatrice  
TITLE OF INVENTION: Peptides and Compounds That Bind  
TITLE OF INVENTION: Selectins Including Endothelial Leukocyte Adhesion  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: 699 Prince Street  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22313  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/241,054  
FILING DATE: 11-MAY-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/057,295

FILING DATE: 05-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/881,395  
FILING DATE: 06-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Gerald F. Swiss  
REGISTRATION NUMBER: 30,113  
REFERENCE/DOCKET NUMBER: 000324-002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-854-7400  
TELEFAX: 415-854-8275  
INFORMATION FOR SEQ ID NO: 85:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-241-054-85

Query Match 64.4%; Score 38; DB 1; Length 20;  
Best Local Similarity 62.5%; Pred. No. 4.3;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TWGQYMOV 9  
DB 7 TWGQWMDL 14

RESULT 28  
US-08-241-054-91  
Sequence 91, Application US/08241054  
Patent No. 5643873  
GENERAL INFORMATION:  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Cwirla, Steven E.  
APPLICANT: Dower, William J.  
APPLICANT: Kollier, Kerry J.  
APPLICANT: Lee, Jung  
APPLICANT: Martens, Christine L.  
APPLICANT: Rubland-Fritsch, Beatrice  
TITLE OF INVENTION: Peptides and Compounds That Bind  
TITLE OF INVENTION: Selectins Including Endothelium Leukocyte Adhesion  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: 699 Prince Street  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22313  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/241,054  
FILING DATE: 11-MAY-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/057,295  
FILING DATE: 05-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/881,395  
FILING DATE: 06-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Gerald F. Swiss  
REGISTRATION NUMBER: 30,113  
REFERENCE/DOCKET NUMBER: 000324-002  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-854-7400  
TELEFAX: 415-854-8275  
INFORMATION FOR SEQ ID NO: 91:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-241-054-91

Query Match 64.4%; Score 38; DB 1; Length 20;  
Best Local Similarity 62.5%; Pred. No. 4.3;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TWGQYMOV 9  
DB 7 TWGQWMDL 14

RESULT 29  
US-08-390-156A-40  
Sequence 40, Application US/08390156A  
Patent No. 5648458  
GENERAL INFORMATION:  
APPLICANT: Cwirla, Steven E.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Dower, William J.  
APPLICANT: Martens, Christine L.  
TITLE OF INVENTION: Peptides and Compounds That Bind to  
TITLE OF INVENTION: ELAM-1  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Affymax Technologies, N.V.  
STREET: 4001 Miranda Ave.  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/390,156A  
FILING DATE: 16-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/057,295  
FILING DATE: 05-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/881,395  
FILING DATE: 06-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Swiss, Gerald F.  
REGISTRATION NUMBER: 30,113  
REFERENCE/DOCKET NUMBER: 1023.1A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-496-2300  
TELEFAX: 415-424-0832  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-390-156A-40

Query Match 64.4%; Score 38; DB 1; Length 20;  
Best Local Similarity 62.5%; Pred. No. 4.3;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TWGQYMOV 9

Db 7 TWGQIMDL 14

# RESULT 30

US-08-390-156A-43

; Sequence 43, Application US/08390156A

; Patent No. 5648458

; GENERAL INFORMATION:

; APPLICANT: Cwifia, Steven E.

; APPLICANT: Barrett, Ronald W.

; APPLICANT: Dower, William J.

; APPLICANT: Martens, Christine L.

; TITLE OF INVENTION: Peptides and Compounds That Bind to

; TITLE OF INVENTION: ELAM-1

; NUMBER OF SEQUENCES: 113

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Affymax Technologies, N.V.

; STREET: 4001 Miranda Ave.

; CITY: Palo Alto

; STATE: California

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/390,156A

; FILING DATE: 16-FEB-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/057,295

; FILING DATE: 05-MAY-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/881,395

; FILING DATE: 06-MAY-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Swiss, Gerald F.

; REGISTRATION NUMBER: 30,113

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-424-0832

; TELEFAX: 415-424-0832

; INFORMATION FOR SEQ ID NO: 43:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 20 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-08-390-156A-43

OY 2 TWGQYMOV 9

Db 7 TWGQIMDL 14

Search completed: July 3, 2001, 10:22:35  
Job time: 1300 sec

